DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

"Beyond the walls:
A mixed methods study of teenagers' extramural English practices and their vocabulary knowledge"

verfasst von / submitted by

Mag.phil. Marlene Schwarz

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktorin der Philosophie (Dr. phil)

Wien, 2020 / Vienna 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 792 343
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	English and American Studies
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Ute Smit

Acknowledgements

Abstract

These days learners of English are exposed to a considerable amount of L2 input during their leisure time. Even in countries where English has conventionally been designated a 'foreign language', it has gained entrance into many areas of our daily lives due to its importance as a global lingua franca. Research on informal language learning outside the classroom is therefore beginning to emerge as a new international research field. In Europe, empirical work has, however, concentrated strongly on countries in which subtitled television programmes constitute a major source of English input, whereas there are comparably few studies on countries where dubbing is the common practice.

This study explores the extramural English (EE) practices of teenagers in one such environment and presents the first larger-scale effort to investigate the relation between engagement with extramural English and vocabulary knowledge among secondary school students in Vienna, Austria. 201 learners attending 10th grade in academic secondary schools participated in this fully integrated mixed-methods study. In the quantitative strand, data on the frequency and amount of participants' out-of-school engagement with English were collected with the help of a detailed questionnaire and a structured online language diary, while their receptive and productive vocabulary size was measured using two vocabulary tests. In addition, focus group interviews were carried out with six groups of learners (N = 30) in the subsequent qualitative strand to gain an understanding of the teenagers' perspectives as they are arguably the most important stakeholders in processes of informal language learning.

Results show that Viennese teenagers' EE environments are characterized by a few very common activities but at the same time they are highly diverse and individualized. The vast majority of participants engage in English activities on a daily basis and on average they spend four hours a day with English during their leisure time. In relation to vocabulary knowledge, regression models indicate a positive relationship between extramural English and receptive, but not productive, vocabulary size, although additional analyses suggest that at least some lexical learning is taking place on a productive level as well. Concerning students' perception of extramural English and informal language learning, the interview data show that participants evaluate engagement with EE as beneficial to their language and vocabulary acquisition although they regard past English teaching at school as the basis for their out-of-school English activities. In comparison to previous studies, particularly those in subtitling countries, the findings of this project thus highlight two important conclusions: the difference between subtitling and non-subtitling countries is negligible in relation to types and amount of contact with extramural English among adolescents, but early exposure to English in subtitling countries appears to have a significant impact on the trajectory and outcomes of informal language learning.

Overall, the study makes a significant contribution to the emerging research field of informal language learning by closely comparing different conceptualizations of the object of study, by introducing methodological innovations with regard to mixing methods and vocabulary measurement, and by empirically exploring a new research environment from several different perspectives.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

BNC	British National Corpus
CATSS	Computer Adaptive Test of Size and Strength
CEFR	Common European Framework of Reference for languages
CLIL	Content and Language Integrated Learning
COCA	Corpus of Contemporary American English
EE	Extramural English
EEOLD	Extramural English Online Language Diary
EEQ	Extramural English Questionnaire
EFL	English as a foreign language
ELT	English language teaching
ESL	English as a second language
ESP	English for specific purposes
FA	False alarm (in a Yes/No test)
FL	Foreign language
IDLE	Informal Digital Learning of English
L1	First language
L2	Second language (including all languages learned after the L1)
LA	Language awareness
LL	Linguistic Landscape
LVLT	Listening vocabulary levels test
MMORPG	Massively Multiplayer Online Role-Playing Game
MMR	Mixed Methods Research
NS	Native Speaker
OILE	Online Informal Learning of English
PVLT	Productive Vocabulary Levels Test
PPVT	Peabody Picture Vocabulary Test
PVST	Picture Vocabulary Size Test
SES	Socioeconomic status
SLA	Second language acquisition (including foreign language acquisition)
VKS	Vocabulary Knowledge Scale
VLS	Vocabulary Learning Strategies
VLT	Vocabulary Levels Test
VOIP	Voice over IP (e.g. Skype)
VST	Vocabulary Size Test

Statistical symbols and abbreviations

α	significance level
[ ]	95% confidence interval of a given point estimate [lower limit, upper limit]
B	unstandardized regression coefficient
β	standardized regression coefficient
CI	95% confidence interval
H	test statistic of the Kruskal-Wallis test
M	mean
Mdn	median
Min	minimum
Max	maximum
p	p-value of statistical significance
SD	standard deviation
r	Pearson's correlation coefficient
r_s	Spearman's rho (Spearman's rank order correlation coefficient)
τ	Kendall's tau
t	t statistic
R^2	r squared (variance explained)
sr^2	squared semi-partial correlation
VIF	Variance inflation factor (for regression models)
W	W statistic (of the Wilcoxon rank-sum test)

1 Introduction

Language pervades all areas of life: languages are used for multiple communicative purposes in many diverse contexts and, equally as important, they are learned in as many different contexts. While this premise is uncontroversial for first language (L1) acquisition, there is a growing realization that it also holds true for all additional languages. In the past, four primary contexts for second language (L2) learning have been identified (Llanes 2018, see also Freed, Segalowitz & Dewey 2004; Munoz 2008): the first is a naturalistic setting, in which the L2 is the main language, or one of several main languages, and in which the learner is continuously surrounded by their second language in diverse situations for an extended period of time. The second context is a foreign language (FL) instructed setting, in which learners acquire an additional language in their home country, or in any other environment in which the L2 does not have an official or historical role, through classroom instruction and have little L2 exposure outside formal education. The third context refers to immersion, which attempts to recreate some aspects of naturalistic settings in FL environments and uses the L2 as a medium of instruction in content subjects. The fourth and final context is study abroad, which shares the characteristics of naturalistic settings, but is limited to a short period of time and frequently connected to prior learning in instructed FL environments.

However, these four contexts do not adequately describe all settings for second language acquisition (SLA) any more. Technological advancements and novel social practices related to these have led to the emergence of a new context which is limited by geographical position to a much lesser extent: extramural engagement. *Extramural* engagement refers to all L2 contact outside the walls of educational institutions and therefore encompasses all forms of L2 activities beyond the walls of classrooms, schools and other formal language teaching environments (Sundqvist 2009a; Sundqvist & Sylvén 2016). The extramural context is linked to the other settings because it is defined in opposition to formal, instructed FL environments, including immersion, and can potentially share some characteristics of naturalistic settings. It constitutes a new hybrid context made possible by the ongoing process of globalization, in which the L2 is used for leisure activities, in particular those that are carried out online and can thus be accessed from anywhere in the world.[1] Hence, extramural engagement presents an additional contextual layer that can potentially be combined with any of the other four contexts for second language learning.

It has already been pointed out that the emergence of this new context is related to globalization and as such it is frequently connected to one particular language. English currently is the globally dominating lingua franca (Crystal 2003; Graddol 2006; Seidlhofer 2001, 2011) and as such it is "an inevitable presence" in many contexts according to Mair (2020: 13):

---

[1] For a more general discussion on the impact of globalization on language learning, see for instance Lo Bianco (2014).

> It [i.e. English] has become the global lingua franca in a comprehensively empirical sense – in international diplomacy and law, science, trade and commerce, tourism, media and pop culture. As such, it is an inevitable presence everywhere, regardless of whether a particular region of the world has a shared colonial history with Britain or, in some cases (e.g. Philippines, Puerto Rico) with the United States.

Due to its position as a global language and its extensive use in media productions and online platforms, there now are numerous possibilities to engage with *extramural English* (EE, Sundqvist 2009a) in contexts where exposure was limited in the past. In fact, it is in relation to English that this new context for language learning first emerged (e.g. Berns, De Bot & Hasebrink 2007; Lamb 2002; Sylvén 2004/2010). At the moment, English is in a unique position with regard to extramural language use, which is, however, not to say that extramural activities in other L2s are not possible or equally valuable; it only indicates that the current conditions and social practices make these a much less widespread phenomenon.

The focus of this thesis is on the extramural English practices of teenagers in Vienna, Austria and their relationship to vocabulary knowledge. The following section provides more detailed information on the rationale for the empirical study as well as its aims and section 1.2 gives an overview of the structure of this thesis and defines key terminology.

1.1 Rationale and aims of the project

Extramural English and informal practices of language learning and use have attracted growing attention over the last decade and by now there is a small body of research on EE and its relation to language development. In Europe, most work in this field has, however, concentrated strongly on countries like Belgium, Sweden or the Netherlands, in which young learners are exposed to large amounts of English input already at an early age through subtitled television programmes. In contrast, children in non-subtitling countries that use dubbing or voice-over strategies for broadcasting foreign language series and films do not come in daily contact with English from an early age onwards. Moreover, large-scale quantitative research conducted for the European Commission (Media Consulting Group/EACEA 2009) indicates that subtitling has an effect on foreign language knowledge: participants in countries with a tradition of subtitling had higher proficiency levels in their additional languages, and particularly in English, than those in non-subtitling countries (see also Rupérez Micola, Bris & Banal-Estañol 2009). For these reasons, calls have been made (e.g. Sockett 2014) for more research on engagement with EE in non-subtitling countries. Austria is such a context because foreign language TV programmes are commonly dubbed into German and thus presents an interesting environment for researching EE complementary to previous studies.

In Austria, English has traditionally been considered a foreign language, although it has gained entrance into many areas of daily life and is now increasingly used in business, education, science, advertising and some media (Archan 2006; De Cillia & Haller 2013; Nagel et al. 2012; Smit & Schwarz 2020, see also Hoffmann 2000). Indeed, a report on the results of the educational standards tests for English in 2013 posits that English is helpful and *necessary* in many situations

of professional and private everyday life (Schreiner & Breit 2014: 22, my emphasis). In this respect, it is questionable whether English in Austria can still be called a 'foreign language' (Smit & Schwarz 2020, see also Gnutzmann & Intemann 2005), but that is not our focus here. Rather, it is on Austrian, or more specifically Viennese, adolescents' EE practices and their relation to language learning in the form of vocabulary acquisition. The empirical study presented in this thesis is the first larger-scale project to systematically explore the link between engagement with EE and language gains in the Austrian context.[2]

Vocabulary knowledge has been chosen as a more specific focus because it constitutes an important component of language competence. It was not feasible to adequately measure overall English proficiency in this study, but receptive vocabulary size is often used as a proxy for general proficiency because it explains large amounts of variance of up to 50% in measures of other language skills (Alderson 2005; Miralpeix & Muñoz 2018). More importantly, results on the relation between EE and vocabulary knowledge among Viennese secondary school students can be usefully compared to previous research in subtitling countries with comparable groups of learners (e.g. Peters 2018; Peters et al. 2019; Sundqvist 2009a; Sylvén 2004/2010) to explore effects of early exposure to English-language TV. The importance of research on the relation between lexical acquisition and extramural English has recently also been highlighted by Schmitt (2019: 267), who calls on researchers to "[d]etermine how to maximize the benefits of extramural exposure for vocabulary acquisition." However, before strategies to support vocabulary learning from EE can be developed, further data are needed.

In addition to investigating the relation between EE and vocabulary knowledge, the study centrally includes learners' views as they are arguably the most important agents in processes of informal language learning and use. Learners' emic perspectives on EE and its potential for language learning have not been considered in many studies so far (e.g. Anioł 2011; Grau 2009; Ingvarsdóttir & Jóhannsdóttir 2017; Lai 2015) and especially previous research on EE and vocabulary has not yet taken them into account.

The empirical study therefore pursues three main objectives: first, it aims to map the landscape of extramural English in Vienna, Austria by investigating the types and amount of contact with English outside formal education among upper secondary school students. Second, it aims to explore the relationship between engagement with EE and both receptive and productive vocabulary size and third, it aims to provide insights into teenagers' perceptions of the phenomenon of EE, its potential for language learning and previous (vocabulary) learning experiences.

---

[2] The second Austrian study on EE and vocabulary knowledge by Hahn (2017) is actually a partial replication of this study and is discussed in more detail in Chapters 3 and 4.

To accomplish these aims, both quantitative and qualitative data collection procedures are integrated in a mixed methods study. The research design consists of a larger quantitative strand using a questionnaire, language diary and two vocabulary tests to gather data on participants' contact with EE and their vocabulary knowledge and a sequential qualitative strand which uses focus group interviews to explore participants' perspectives in greater detail. The integration of both quantitative and qualitative data in this exploratory study thus allows to gain a more in-depth understanding of extramural English and its relation to vocabulary knowledge in a relatively new research environment.

1.2 Outline

Following this brief introduction to the empirical study presented in this thesis, Chapter 2 describes the concept of extramural English more closely and puts it in relation to other relevant conceptual approaches to the phenomenon of informal learning through leisure activities, such as *language learning beyond the classroom*. In addition, it provides an overview of the small existent body of research through an extensive literature review.

Chapter 3 is concerned with the second area of interest in this project and discusses the foundations of vocabulary research and measurement as well as the current state of knowledge about L2 vocabulary development as a background to the methodology and results of this study. It then zooms in more closely on the focus of the empirical study by synthesizing research on vocabulary learning from EE and from specific activities such as reading, listening, viewing or gaming.

Chapter 4 provides more specific information on the research context: after a brief introduction to the linguistic situation in Austria, the roles of English in education and the public sphere are explored, the latter being exemplified by English in the linguistic landscape, in business and in the media. In addition, this chapter also contains more general information on Austrian adolescents' lifeworlds, such as their leisure activities and media use.

Chapter 5 introduces the empirical study: it presents the research questions and the study design after outlining the principles of a mixed methods approach. Subsequently, it sets out the details of the research design with regard to sampling and participants, research instruments, data collection procedures, and methods of analysis for both the quantitative and the qualitative strand.

The results of this study are presented in the following two chapters: Chapter 6 presents the results of the quantitative strand in relation to participants' background, their use of extramural English and its relation to receptive and productive vocabulary knowledge. In parallel, Chapter 7 reports the findings of the qualitative strand, which provide insights into the significance of English in participants' everyday lives, their EE practices, learning from EE with a focus on vocabulary acquisition, and the relationship between in- and out-of-school uses of English.

Chapter 8 first presents the conceptual and methodological contributions of this thesis before drawing together the results of the quantitative and the qualitative strands and discussing them in light of the literature presented in Chapters 2, 3 and 4. Concluding this thesis, Chapter 9 highlights key findings and their implications before addressing the significance of the study as well as its limitations and directions for further research.

Before going in medias res, some terminological issues still need to be clarified. In the remainder of this thesis I will use the term *second language* or *L2* to refer to all languages learned after and in addition to one or more first languages. Although this term has recently been criticized by Dewaele (2018), who suggests the use of *LX* as a more value-neutral label, such an inclusive use of the term L2 is in line with suggestions in the SLA literature (e.g. Doughty & Long 2003; Ellis 1994, 2015; Mitchell, Myles & Marsden 2013; Ortega 2011). Most notably it is used in an article proposing "A transdisciplinary framework for SLA in a multilingual world" by The Douglas Fir Group (2016), a group of eminent scholars based in North America, who state that

> we define the object of inquiry of SLA as additional language learning at any point in the life span after the learning of one or more languages has taken place in the context of primary socialization in the family; in most societies this means prior to formal schooling and sometimes in the absence of literacy mediation (The Douglas Fir Group 2016: 21, emphasis in original).

In keeping with the use of L2, the term *second language acquisition (SLA)* is used as an umbrella term for the acquisition of all additional languages including what has conventionally been called foreign language learning. In addition, I do not differentiate between the terms language acquisition, learning and development, these are used synonymously throughout the text. Finally, L2 learning is seen as encompassing both cognitive and social processes and being shaped by diverse social conditions, cognitive resources and individual experiences. Such a view has again been succinctly summarized by The Douglas Fir Group (2016: 36):

> Language learning is a complex, ongoing, multifaceted phenomenon that involves the dynamic and variable interplay among a range of individual neurobiological mechanisms and cognitive capacities and L2 learners' diverse experiences in their multilingual worlds occurring over their life spans and along three interrelated levels of social activity: the micro level of social action and interaction, the meso level of sociocultural institutions and communities, and the macro level of ideological structures.

2   Extramural English and language learning beyond the classroom

Learning in informal contexts through leisure time pursuits has received increased attention from several research fields and discourses in recent years. While the area of interest in this project is more defined narrowly as language learning beyond the classroom and extramural English (EE), it is important to recognize the various discourses that have recently contributed to discussions of informal learning. Owing to different research perspectives and foci, different topics and terms have been in the foreground of these discourses (see also Blell 2015; Werquin 2016), which clearly shows that the emerging research field concerned with learning (and teaching) outside formal educational contexts lies at the intersection of several research strands. Currently, these strands often only overlap to a very limited extent, although greater integration could inform theoretical conceptualizations and promote the clarification of terminology within related discourses. While a comprehensive overview of the various discourses contributing to the field of learning beyond formal education is well beyond the scope of this study, it is useful to consider pertinent research strands that could be relevant for future language learning related research.

Unlike the Anglophone context, which has a much longer tradition of theorizing and researching non-formal education and informal learning (for historical overviews see Overwien 2005, 2016; Rohs 2016a), in Europe interest in informal learning has grown mainly as a result of the European Commission's (2000) memorandum on lifelong learning, which recognizes and promotes formal, non-formal and informal learning as part of lifelong learning (see also Colley, Hodkinson & Malcom 2003). As a result, the European Centre for the Development of Vocational Training (CEDEFOP) has focused on the role of learning in the workplace and the validation of non-formal and informal learning (e.g. CEDEFOP 2011, 2015), while other research reports and reviews on informal learning have been intended to inform educational policy (e.g. Dohmen 2001 in Germany, Sefton-Green 2004 in the UK). It is important to note that pedagogical discourses on informal learning (see the edited volumes by Harring, Witte & Burger 2016; Rohs 2016b) have focused on various subjects, ranging from science (e.g. Bell et al. 2009, Uitto et al. 2006) over history (Karpa, Overwien & Plessow 2015) to literacy (Rogers 2008) and media literacy (Livingstone 2014). Similarly, different physical locations and contexts, such as museums (Crowley, Pierroux & Knutson 2014; Rymarczyk 2015), historical sites (Klein 2015), volunteer work (Düx & Sass 2005) and even airports (Legutke & Thiel 1983) have been identified as spaces for out-of-school learning. Especially in the German-speaking academic community, there is a broad discourse on *außerschulische Lernorte* [learning spaces outside school] (see Burwitz-Melzer, Königs & Riemer 2015, Karpa, Overwien & Plessow 2015, Rohs 2010), which, however, rarely intersects with international (English-language) research. Recently, digital learning spaces have attracted attention and research on informal learning seems to concentrate increasingly on the affordances offered by the latest technological advances and digital learning

in online contexts (e.g. Drotner 2008; Ito et al. 2010; Lemke et al. 2015; Sefton-Green 2013). While some of these developments also hold true for language-related research on informal learning outside school, particularly the more recent focus on learning in online environments, it appears that overall the different research strands have been developing largely independently of each other. I would argue that a greater awareness of the developments in other areas of the larger research field could positively influence research on informal learning outside educational institutions, particularly in terms of theoretical conceptualizations, and lead to unexpected synergistic effects.

In the following, this chapter will concentrate on language learning beyond the classroom and more specifically on learning English through extramural practices. First, an endeavour to model language learning beyond the classroom is presented (section 2.1) before describing and delineating the most central concept for this study: extramural English (section 2.2). The last part of this chapter (section 2.3) provides an overview of empirical research on language learning in extramural contexts.

2.1 Language learning beyond the classroom

Out-of-school language learning is not a new phenomenon, but "while opportunities for learning beyond the classroom have always been available to learners, technology and the Internet have dramatically expanded both the scope and nature of these opportunities" (Nunan & Richards 2015b: xii). In line with the increasing opportunities for learning languages, and in particular English, outside the classroom, interest in such learning situations is growing exponentially among L2 researchers. Benson and Reinders (2011b: 5) identify the social turn in SLA research as another driving force for such research, stating that learning is no longer seen "in purely cognitive terms, but in terms of participation in communities and contexts of various kinds". They further argue that the increasing number of publications on *language learning beyond the classroom* indicates the emergence of a new "area of inquiry with its own theoretical assumptions and issues of particular concern" (Benson & Reinders 2011b: 5).

In the same way that language learning outside the classroom is not a new and unprecedented phenomenon (see the review in section 2.3), research on such language learning is not a novel endeavour either. The introduction to this chapter mentioned the wealth of terms and concepts used in the wider area of informal learning, and the emerging research field of language learning beyond the classroom is very similar in this respect. Different authors use different terms to describe their object of study, such as out-of-class learning (e.g. Benson 2001; Cole & Vanderplank 2016; Verspoor, De Bot & Van Rein 2011; Yap 1998), informal language learning (e.g. Lee & Dressman 2018; Sockett 2014), incidental language learning (e.g. De Wilde & Eyckmans 2017; Kuppens 2010) or extramural English (e.g. Sundqvist 2009a; Olsson 2016). As Benson and Reinders (2011b: 1) point out, the area of enquiry is frequently defined by what it is *not*, using terms like *out*-of-school or *out*-of-class learning, *in*formal or *non*-formal learning, or *non*-instructed learning. Benson and Reinders (2011b) thus propose *language learning beyond*

the classroom (LBC) as a new and more inclusive term that reduces the risks of defining the object of enquiry exclusively in negative terms and of "treating the world beyond the classroom as an alternative to the classroom, as if classrooms were the natural place for language learning to take place and the world beyond the classroom a strange and hostile territory in which languages are learned with difficulty, if at all" (Benson & Reinders 2011b: 1). In an attempt to further reduce terminological confusion and to define the scope of the emerging research field Benson (2011) suggests a preliminary model of LBC. He shows that the various terms used to describe such learning actually "point to those four distinct dimensions of language learning beyond the classroom – location, formality, pedagogy, locus of control" (Benson 2011: 9). Figure 2.1 presents a visualization of Benson's (2011) model showing how the four dimensions interrelate to characterize any given learning situation.
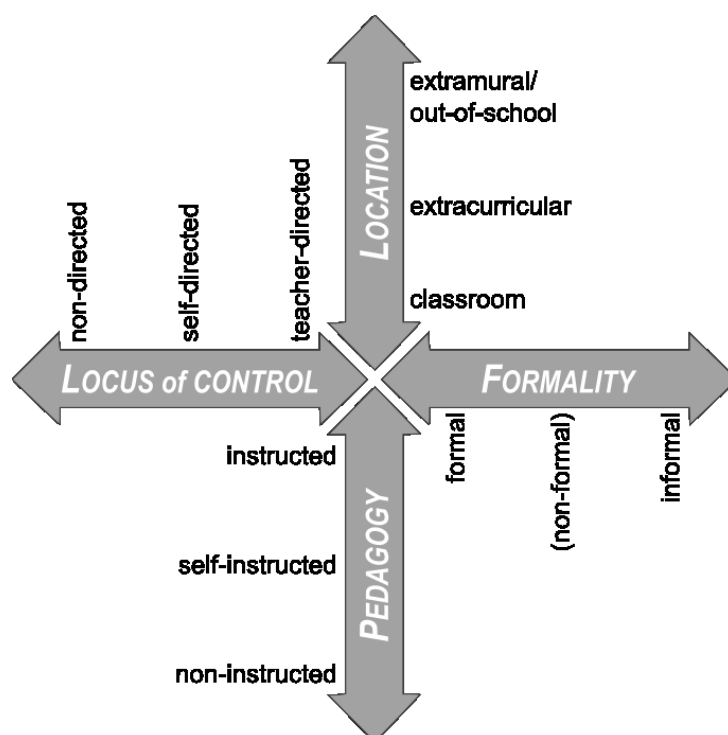
Figure 2.1: Graphic representation of language learning beyond the classroom based on Benson (2011)

The dimension of *location* is perhaps the most tangible, it simply refers to the setting for learning. Points on the continuum of location indicate different distances from the instructed environment of the language classroom, with extracurricular or after-school activities still happening at school and extramural or out-of-school learning taking place at different sites, such as the learners' homes or places for hobbies. It is important to note that LBC also includes private school lessons or individual tutoring session, since these are located outside the official school context, even though they may be comparable to the language classroom in terms of formal and pedagogical aspects. Another point worth mentioning is that Benson (2011: 9) uses the term extramural as a synonym for extracurricular and thus suggests that extramural activities still take place at school, though outside regular lessons. In line with Sundqvist's concept of extramural English (see section 2.2), I use extramural as equivalent to out-of-school in the

context of this study; therefore, the term has been placed at the far end of the continuum of the location dimension in Figure 2.1.

The next dimension, *formality* is related to educational institutions and certification, thus formal learning situations involve teaching and assessment, while informal ones typically do not. Benson (2011: 10) explains that "[t]he dimension of *formality*, therefore, essentially refers to the degree to which learning is independent of organized courses leading to formal qualifications". He also makes a distinction between non-formal and informal learning stating that "non-formal education often refers to classroom or school-based programmes that are taken for interest and do not involve tests or qualifications, while informal education refers more to non-institutional programmes or individual learning projects". This definition of non-formal and informal learning is, however, debatable since in their review on informality and formality in learning Colley, Hodkinson and Malcom (2003: vi, executive summary) come to the conclusion that "[t]here is no clear difference between informal and non-formal learning. The terms are used interchangeably, with different writers expressing preferences for each". Instead, the authors suggest that "[i]n practice, elements of both formality and informality can be discerned in most, if not all, actual learning situations [...]. In other words, formality and informality are not discrete types of learning, but represent attributes of it" (Colley, Hodkinson & Malcom 2003: executive summary). The perspective that attributes of both informality and formality are present in all learning situations, is wholly compatible with the conceptualization of formality as a continuum, if it is assumed that the extreme endpoints of completely formal or informal learning do not exist in reality. In contrast, non-formal learning does not fit this conceptualization well; hence, it has been included in the graphic representation of Benson's model in Figure 2.1 for the sake of completeness, but has been put in parentheses for the reasons given above.

The third dimension of *pedagogy* can range from instructed learning over self-instructed learning to non-instructed or naturalistic learning. Benson (2011: 11) sees instruction as a "particular kind of pedagogy, involving formal processes such as sequencing of material, explicit explanation and testing". Instruction is typically teacher-led and occurs inside the language classroom, while self-instruction is learner-led and "specially designed books or television or radio broadcasts take on the role of a classroom instructor and there is a strong intention to learn on the part of the learner" (Benson 2011: 11).[3] In non-instructed, or as Benson calls it, naturalistic language learning "there is no instruction [...] and, in principle, no intention to learn" (Benson 2011: 11); thus, learning from language activities that are primarily carried out for entertainment purposes could be considered non-instructed learning. Although these terms have been placed on a continuum in Figure 2.1 above, it is worth adding that a classroom setting does not per se imply that instruction is taking place (see also Benson 2011: 11), since other forms

---

[3] While the role of instructor in classroom settings is most typically taken on by teachers, that is not to say that other actors such as peers in and out of the classroom or parents cannot perform the role of instructor.

of learning like exploratory learning or self-instructed learning do and should play a role within institutional settings as well.

Finally, language learning situations can also be differentiated in terms of who holds control over learning. *Locus of control* can lie with an instructor as in teacher-directed learning or with the learner as in self-directed, autonomous or independent learning, or, I would add, it can even be non-directed, when no conscious decision to learn has been made.[4] To give an example, a home tutoring lesson could be characterized as an instructed, relatively formal learning situation in an out-of-school setting, in which the tutor holds most control over the decisions for learning. In contrast, looking up an unknown word a learner heard in a pop song is a typical informal situation in an extramural context. Depending on the goal of the learner, this situation can be regarded as involving a high degree of autonomy or self-direction and as being self-instructed if the aim is to learn a new word, or it could be seen as a non-directed and non-instructed learning situation if the learner's purpose simply is to understand the song. Hence, as Benson (2011: 12) concludes there "appears to be no simple relationship between the location of learning (in or out of class) and locus of control". The examples given above also indicate that all dimensions are present in any given learning situation and can interact in many different ways (see also Benson 2011: 12). Benson (2011: 13) concludes that

> [o]ne way of defining language learning beyond the classroom as a field of inquiry, therefore, is to say that it is centrally concerned with locations for language learning other than the classroom and with relationships between these locations and aspects of formality, pedagogy and locus of control.

Consequently, he argues that in terms of research on such language learning situations, we need to look at both the *settings* and the *modes of practice*, the forms of learning and teaching that take place in such settings. Benson (2011: 13) regards settings as more than physical locations, they are described as "social spaces" or "a particular set of circumstances within a location that offer affordances for and constraints on the possibilities for language learning". Setting thus includes the social configuration and the pedagogical relationships between the actors as well as the availability of both physical and virtual resources and materials. The learning activities which take place in such a setting are the modes of practice, which Benson (2011: 14) defines as "a set of routine pedagogical processes that deploy features of a particular setting and may be characteristic of it".[5] According to Benson (2011: 14) settings and modes of practice are "relatively independent" of each other, which means that first, a setting can support more than one mode of practice and second, that the modes of practice learners actually make use of in a given setting cannot be predicted because individual learners perceive and act upon different affordances in a given context (see the example given in Benson 2011: 14–15).

---

[4] For this reason "non-directed learning" was added as the endpoint of the locus of control continuum in Figure 2.1, although Benson (2011: 12) does not mention it.
[5] Here, it is important to highlight again that pedagogy is not the same as instruction, which is only one particular type of pedagogy in Benson's (2011) view. He uses pedagogy in a broad sense, whereas instruction is defined more narrowly as "knowledge transmission" (Benson 2011: 16).

Although Benson's definition of modes of practice as "routine *pedagogical* processes" understandably stems from an interest in *learning* beyond the classroom, it begs the question how, in practice, researchers are meant to distinguish such routine pedagogical processes from other social practices. Defining modes of practice as pedagogical processes implies that there is an intention to learn or a learning aim on the part of the learner, but in empirical research such an intention is difficult to establish and/or trace. It is therefore virtually impossible to investigate only routine *pedagogical* processes because learners engage in many language practices outside the classroom and we cannot know a priori which of these lead to learning. Consequently, a more inclusive conceptualization of modes of practices as routine social practices which involve target language use and have pedagogical potential may be more practical. I therefore suggest extending Benson's definition of modes of practice to better reflect the diverse processes learners engage in:

> A mode of practice is a set of routine social practices which are located in and deploy features of a particular setting, involve target language use and have pedagogical potential.

According to this definition a mode of practice includes both intentional and incidental learning and does not make any assumptions about the learner's aim or purpose. To give an example, a mode of practice set in an Austrian learner's home on a normal weeknight could be watching an English-language series with subtitles. The learner might engage in this activity intentionally to immerse themself in the language in preparation for an English exam the following day or they might watch the series in English because the dubbed German version is not yet available. Regardless of the learner's intention, watching the series in English could improve their listening comprehension, help them realize a pronunciation error or lead to the acquisition of a new lexical item. However, whether the pedagogical potential of this mode of practice will be realized, even if the learner fully intends to do so, is impossible to tell in advance, which underlines the problem of operationalizing Benson's (2011) definition for empirical research and adds support to the more inclusive working definition suggested above.

While certainly a highly useful first attempt at establishing a framework for describing and researching language learning beyond the classroom, Reinders and Benson (2017) emphasize that Benson's (2011) model is rudimentary and could be extended to include several other dimensions not currently considered in the theoretical framework. Indeed, a few researchers who have taken up Benson's model have made suggestions for additional aspects to be included in the model: Chik (2014) uses the framework to analyse interview data from a project on L2 gaming and learning practices and found that she needed an additional dimension describing the development of and interrelation between gaming and learning over time, which she called *trajectory*. In Lai, Zhu & Gong (2015) the model forms an important backdrop for their study on the out-of-class learning experiences of middle school students, but results suggest that the "degree of diversity in the overall language learning ecology could be an appropriate criterion to evaluate the quality of out-of-class" (Lai, Zhu & Gong 2015: 298). Thus, Reinders & Benson (2017) suggest that *variety of activities* could be yet another component of an improved model of

LBC. They list several further dimensions that could be added to develop the framework further including *mediation* in relation to resources used, *sociality* in relation to the social situatedness of learning processes, *modality* in relation to the modes of language study or use, a *linguistic* dimension in relation to skills and levels of competence, and dimensions focusing on the *intentionality*, *explicitness* or *inductiveness* of language learning processes (Reinders & Benson 2017: 562–563, for further suggestions see also Kurtz 2015: 109–111; Richards 2015: 19). While all of these are potentially very relevant concepts in the context of LBC, a model including ten or more dimensions may become impractical in terms of operationalizations for research. In addition, information on some of these aspects like sociality, mediation or modality can, and according to Benson's (2011) definition should be, provided in the descriptions of settings. Hence, in practice individual researchers are likely to adapt and add to the model in accordance with their research context and focus following the example of Chik (2014).

In addition to language learning beyond the classroom several other researchers in the emerging research field have proposed their own terms and concepts to reflect their particular research focus. In general, these tend to be more specific than Benson and Reinders' (2011a) rather inclusive approach. As we shall see in the next section, Sundqvist (2009a: 26) sees her concept of *extramural English* (EE), which is central to this study, as an umbrella term, although LBC potentially includes an even broader spectrum of learning situations than EE. The relationship between these two concepts, which represent complementary perspectives, is discussed in more detail in section 2.2. In contrast, the concepts of *online informal learning of English* (OILE) proposed by Sockett (2013, 2014) and *informal digital learning of English* (IDLE) introduced by Lee (2019a, 2019b; Lee & Dressman 2018) clearly differ from both LBC and EE because they focus exclusively on language learning in digital environments and thus are much more specific.

The main difference between these two concepts seems to be that OILE solely includes activities carried out online, whereas IDLE can potentially also include activities involving digital media in offline contexts. Furthermore, they differ in terms of their perspective on learner's intention to learn when using English. Lee (2019a: 768) states that "IDLE can be conceptualized as self-directed, informal English learning using a range of different digital devices (e.g., smartphones, desktop computers) and resources (e.g., web apps, social media) independent of formal contexts". Lee (2019b) differentiates between extracurricular IDLE, which is other-directed and linked to a formal programme, and extramural IDLE, which is self-directed, naturalistic learning in informal, out-of-class environments. Evidently, both the studies presented in Lee (2019a, 2019b) and the discussion here focus on the latter type, but it is worth noting that Lee (2019a, 2019b) characterizes such language learning as self-directed and thus intentional. In contrast, Sockett (2014: 7) explains that OILE "is best understood as a complex range of internet-based activities", which are mainly carried out for leisure purposes, but also lead to informal language learning. Sockett (2013) thus stresses the incidental nature of such learning and situates it within the context of dynamics systems theory. Consequently, another major difference between the

two concepts is that Lee conceptualizes IDLE as self-directed learning, whereas Sockett sees OILE as an unintentional, incidental by-product of English-language leisure activities carried out online. Both concepts, however, concentrate exclusively on learning with digital media, which happens mainly or entirely online. While a focus on the new affordances of emerging technology is certainly fascinating and understandable in this respect, history has shown that new technologies do not necessarily supersede old ones, but can be added to media environments, depending on individual users' preferences (Hasebrink 2007: 90–92; Johnsson-Smaragdi 2009: 170–171). As will be shown in the following section, extramural English takes both digital and analogue, or 'old' and 'new', media into account and is thus a more fitting choice for the present study.[6]

A further concept to describe instances of language learning outside formal education that has recently emerged is *language learning in the wild* (Eskildsen & Cadierno 2015; Wagner 2015). It is rooted in conversation analytic studies of interactions between L2 learners and members of the target language community, which are not controlled by teachers or researchers and thus occur 'in the wild'. Drawing on a usage-based approach to language learning, previous studies of language learning in the wild (e.g. Eskildsen 2018; Eskildsen & Theodórsdóttir 2017) have investigated "how people make use of everyday encounters with L1 speakers to learn the L2" (Eskildsen & Cadierno 2015: 4). As these studies have exclusively been concerned with second language contexts in Kachru's (1985) original sense up to now, language learning in the wild is not a relevant concept for this study which investigates English language learning in what is conventionally called a foreign language context.

From the above description it should have become clear that the concepts of OILE, IDLE and language learning in the wild present more specific research interests than language learning beyond the classroom. While offering interesting perspectives, they are less relevant to the present study due to their focus on digital or second language environments. LBC and extramural English (see section 2.2) are more appropriate as a theoretical backdrop to a study that seeks to describe Viennese adolescents' English-language environments in out-of-school contexts as comprehensively as possible (see also sections 1.1 and 5.1). More generally, one important question in this developing research area is not only what to call the object of study, but also what exactly to include in it because, as we have seen above, currently the various perspectives differ with regard to crucial details. Clearly, all approaches to informal language learning outside formal educational contexts that are available so far are in need of further development.

One attempt to steer the direction of future research has recently been proposed by Reinders and Benson (2017). They have set out a research agenda for the emerging field of language

---

[6] I agree with Lee's (2019b: 123) note that concepts such as language learning beyond the classroom, EE, IDLE and OILE are not mutually exclusively and share some important characteristics. However, I would argue that at the same time it is worth noting the differences; first, to specify exactly what is the area of investigation in an empirical study, and second, to promote further theoretical reflections within the emerging research field.

learning beyond the classroom, which is partly taken up in the present study. In addition to a need for further developing Benson's (2011) theoretical model (see description above), they propose to focus on three aspects: describing typical settings, analysing learners' experiences, and exploring teachers' thoughts on and support of LBC. In their view, research documenting settings for LBC could involve both situations of study at home or abroad and focus on a particular group of learners or a particular setting. Investigations of learner experiences could concentrate on key experiences and their effects on subsequent language learning, the role of technology in learning processes, or learners' use of strategies. Finally, projects involving teachers' perspectives on LBC could include teacher beliefs about learning outside class, the inclusion of out-of-class learning assignments in language programs, or the provision of learning skills training as a preparation for LBC. In their description of nine research tasks to advance the field, Reinders and Benson (2017) also consider the prerequisites for such projects, including suggestions for useful sequencing of the proposed tasks, and emphasize the importance of teachers taking an active role in this kind of research, for instance through action research projects.

This section described the theoretical framework of language learning beyond the classroom, which constitutes a first and very valuable attempt to disentangle the various terms and concepts used by different researchers and to delimit the scope of the emerging research field. Although still preliminary, Benson's (2011) model is useful to characterize learning situations in and outside instructed language learning contexts and is thus taken up in section 4.4 to describe the specific research context and object of study of this project. Similarly, Benson's suggestion that research on LBC needs to focus on both the settings and the modes of practice actualized within these is certainly beneficial, but from an empirical perspective his conceptualization of modes of practice as routine pedagogical processes is slightly problematic. Therefore, a new way of defining modes of practice which can be implemented more easily in empirical research has been proposed. This new working definition is appropriate for this project because it improves compatibility of the two central theoretical frameworks of extramural English and LBC, but it certainly needs to be adapted and refined to fit other research contexts as well. The next section presents and closely examines the concept of extramural English and further elaborates on its relation to language learning beyond the classroom.

2.2   Extramural English

This section sets out to explain and contextualize the central concept of this study: *extramural English* (EE). Coined by Pia Sundqvist in 2009, it is a compound of Latin origin referring to 'English outside the walls' (Sundqvist 2009a; Sundqvist & Sylvén 2016). More precisely, it refers to 'English outside the walls of educational institutions' and therefore encompasses all forms of informal contact with English outside schools and other instructed language teaching environments. Sundqvist (2009a: 25) provides the following explanation of the scope of her concept:

> In extramural English, no degree of deliberate intention to acquire English is necessary on the part of the learner, even though deliberate intention is by no means excluded from the concept. But what is important is that the learner comes in contact with or is involved in English outside the walls of the English classroom. This contact or involvement may be due to the learner's deliberate (thus conscious) intent to create situations for learning English, but it may equally well be due to any other reason the learner may have. In fact, the learner might not even have a reason for coming in contact with or becoming involved in extramural English.

Hence, EE is defined primarily by its location and does not make any claims about the purpose of engaging with English. Sundqvist (2009a: 26) further adds that "contact with extramural English, or involvement in extramural English activities, is generally voluntary on the part of the learner". Typically, EE activities are informal leisure activities, in which young people engage with English voluntarily during their spare time. Examples include watching movies or TV series, reading posts on blogs or other kinds of websites, interacting with friends on social media platforms, listening to music, playing digital games, or reading articles or books (see also Sundqvist & Sylvén 2016: 7).

Sundqvist and Sylvén (2016: 6) further specify and emphasize the fact that engagement with EE is learner-driven:

> [t]his [EE] contact or involvement is *not* initiated by teachers or other people working in educational institutions; the initiative for contact/involvement lies with the learner himself/herself or, at times, with someone else, such as a friend or a parent. Thus, in general, involvement is voluntary on the part of the learner, though there is the possibility that learners engage in specific EE activities because they feel pressured to do so, for whatever reason (emphasis in original).

To clarify, in my understanding the definitions above imply that EE does not include any homework or other assignments set by teachers, exam preparation, extensive reading programmes, school theatre trips, tutoring sessions or private language courses. Self-access centres for autonomous study are, although mentioned by Sundqvist (2009b), not a typical EE activity either because usually language teachers strongly advise their students to make use of such facilities, if they are provided by the educational institution.

Sundqvist and Sylvén (2016: 10–14) also propose a model of L2 English learning including EE: taking up Benson's (2011: 12) comment in relation to his model of language learning beyond the classroom that there "appears to be no simple relationship between the location of learning (in or out of class) and locus of control" (see section 2.1), they attempt to further theorize this relation. Their model, which is reproduced in Figure 2.2, includes two axes representing the driving force for and the physical location of English learning. These constitute a coordinate system, in which the midpoint represents a fictitious activity that is half learner- and half other-initiated and located exactly at the wall of the English classroom.[7] In addition, in their conceptualization locus of control or driving force, represented by the x-axis, is related to both

---

[7] It is worth bearing in mind that Sundqvist and Sylvén (2016: 12) highlight that the distances on the y-axis (representing the location of learning) below and above the midpoint do not correspond to real-life geographical distances.

the formality and the intentionality of learning with other-initiated activities typically being more formal and intentional than learner-initiated activities, which tend to support informal and incidental learning (Sundqvist & Sylvén 2016: 11–12).
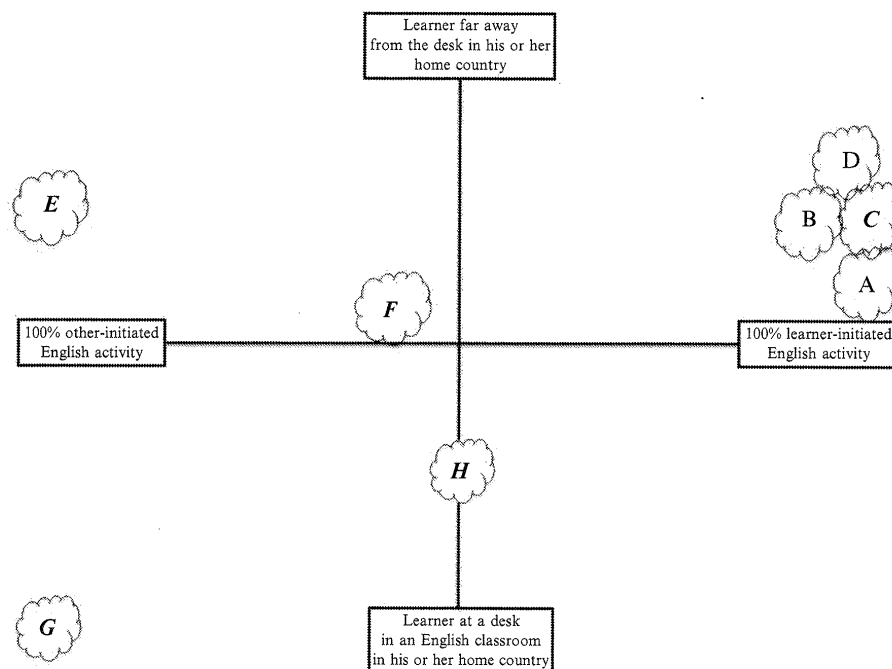
Figure 2.2: Visualization of the model of L2 English learning including extramural English taken from Sundqvist and Sylvén (2016: 10)

The authors then situate different learning activities within the coordinate system, explaining that EE activities are located in the right upper quadrant of the coordinate system as they are learner-initiated and take place outside the walls of the English classroom. In doing so, they show that EE can take place in various contexts like the home (letters B and C), online environments accessed from home (letter D), or even school (letter A).[8] The last point warrants further comment: Sundqvist and Sylvén's (2016) model clearly illustrates that EE activities can take place at school. At first glance, this may seem to contradict its definition in terms of location as 'outside the walls of the English classroom', or as I stated above, as 'outside the walls of educational institutions'. However, schools are not exclusively places for teaching and learning, and although the core element of the EE definition is its location, it is important to recognize that even at school students have periods of leisure such as break times or free periods, in which they can engage in the same EE activities as in their homes, particularly through the use of smartphones, which give them instant access to the virtual world of the internet.

Sundqvist and Sylvén's (2016: 10) model of L2 English learning explicitly includes places where the "learner [is] far away from the desk in his or her home country" (see also Figure 2.2), and thus also study abroad contexts. However, they point out that stays abroad cannot readily be

---

[8] For more detailed explanatory comments on the activities represented by the letters A to H, please see Sundqvist and Sylvén (2016: 12–13).

characterized as one single language activity and therefore can include both classroom-based and extramural practices. To illustrate this issue, they suggest that one could shade a larger area in the coordinate system, if one wanted to represent learning experiences like those in study abroad programmes (Sundqvist & Sylvén 2016: 13). I fully agree that stays abroad can and should not be regarded as one single language activity and would go as far as suggesting that study abroad experiences should not be regarded as part of extramural English, but as a potential context for EE (see also section 1.1). Typically, we are interested in the EE activities in a specific learning context, therefore the learner's language learning experience in the home environment and during the study abroad programme would then constitute two different contexts, both of which include formal and informal language learning and both of which could be explored with regard to EE activities.

Two other crucial points in relation to extramural English are that first, one does not always have to *do* an EE activity and that secondly, EE does not necessarily involve *learning*. With regard to the first point, Sundqvist and Sylvén (2016: 11) state that

> the doing [of an EE activity] does not necessarily imply any action (in its literal sense) on the part of the learner; the doing of an activity could simply mean that a learner is exposed to English (for instance, there is an ad in English in a newspaper).

Hence, EE includes exposure to English by coincidence as well as deliberate use of English for a specific purpose in the same way as it includes receptive and productive language use. Concerning the second point, it is important to highlight that the term *extramural English* does not make any reference to learning, although, clearly, there is an interest to discover whether and how learning occurs through EE activities. However, in contrast to other terms such as *language learning beyond the classroom* (LBC), *online informal learning of English* (OILE) or *informal digital learning of English* (IDLE), EE does not posit learning as a given. EE activities may and often do lead to the acquisition of new language structures, the extension of existing knowledge, or to practice effects for language skills, but they do not always and automatically do so. The research process thus typically involves first charting the EE environment of a particular group of learners or an individual learner and then to see where, when and how learning takes place. This stands in contrast to Benson's (2011) model of LBC (see section 2.1), which is primarily interested in those situations where learning occurs, although these are difficult to perceive and distinguish from situations of language use in advance. Thus, on the one hand, LBC is a broader concept than EE because it includes formal learning and teaching situations in out-of-school contexts (Benson 2011, see also Bailly 2011; Murray 2011 in the same volume), whereas EE focuses on voluntary, learner-driven activities. On the other hand, EE can be considered more inclusive than LBC because it considers all English-language practices a learner engages in outside school, regardless of whether learning actually takes place, whereas LBC is interested specifically in instances of learning in its original conceptualization.

This difference in perspective also reflects the different origins of the two frameworks: while LBC was devised from a theoretical viewpoint to delimit and describe an emerging area of enquiry and to reduce terminological confusion, EE stems from a practical interest in what

learners do and thus reflects a primarily empirical perspective. The new working definition for Benson's concept of mode of practice proposed in section 2.1, which does not presuppose the existence of pedagogical processes, but rather emphasizes the pedagogical potential of extramural language activities, is one attempt to bring the two theoretical concepts closer together so that they can be combined usefully in the present study. Another way is to characterize the object of the empirical study – extramural English practices of Viennese teenagers and their relation to vocabulary knowledge – by locating it on the four dimensions of location, formality, pedagogy and locus of control suggested in Benson's (2011) preliminary model (see section 4.4).

This section provided a detailed explanation of extramural English, the central concept of this study, as conceived by Sundqvist (2009a) and Sundqvist and Sylvén (2016) and clarified its relationship to the second theoretical framework used, language learning beyond the classroom. In this study extramural English practices are defined as English-language activities which take place outside the walls of educational institutions in Austria during learners' leisure time and which are learner-driven and typically voluntary and informal. This working definition thus excludes any activities that are set by teachers or other educational practitioners as well as study abroad experiences because of the reasons mentioned above and the specific interest in EE in Vienna, Austria.

2.3   Extramural English and language learning

Traditionally, SLA studies have largely concentrated on classroom settings, but already in the 1980s researchers studying good language learners (e.g. Rubin & Thompson 1982) showed that these use opportunities to practise the language outside class. For this reason, early studies involving out-of-school language learning often developed from an interest in the strategies of good language learners or from a learner autonomy perspective during the 1990s (e.g. Freeman 1999; Littlewood & Liu 1996; Pickard 1996; Yap 1998). In the early 2000s the developments appear to accelerate, one frequently cited call for research on out-of-class language learning, which could be seen as a juncture in the development of this area of study, was made by Benson in his 2001 book *Teaching and researching autonomy in language learning*. He called on researchers to investigate "out-of-class learning [... as] a new area of study of great importance to the theory and practice of autonomy" (Benson 2001: 203). Benson's call for the development of such a research strand was timely in response to changing glocalized learning environments and this may be one reason why his call has also been taken up outside the field of learner autonomy. Almost 20 years later, this new field of research is rapidly becoming more and more prolific and gaining prominence worldwide, so that by now there is quite a substantial body of research.

In the following, this section aims to provide a brief survey of existing research on extramural language learning by giving an overview from a meta-analytic perspective.[9] The "burgeoning body of literature" (Lai, Hu & Lyu 2018: 115) can be characterized along five dimensions: it can be described in terms of a temporal dimension showing the increasing interest in this field, in terms of location or context, researchers' background and research interests, theoretical concepts and content focus, and the research designs used.

The review includes 77 empirical studies published before 2019, which, despite the use of various theoretical concepts and terms to describe their specific focus (see section 2.1), are united by an interest in language learning taking place outside instructed learning contexts. The studies included in this section investigate informal out-of-school learning environments, or particular aspects relating to these, and examine a range of EE activities rather than focusing on one specific activity. Hence, studies on learning from one particular activity such as gaming (e.g. Chik 2014; Hannibal Jensen 2017; Reinders & Wattana 2010, 2015; Sundqvist & Sylvén 2012, 2014; Sylvén & Sundqvist 2012b) or watching audiovisual media (e.g. Koolstra & Beentjes 1999; Kusyk & Sockett 2012; Lin 2014; Wang 2012) as well as studies focusing on interaction on one particular platform such as Facebook (e.g. Alm 2015), Twitter (e.g. Gleason 2015) or YouTube (e.g. Benson 2015) have not been incorporated for reasons of scope. While a complete overview of all research recently conducted in this area is virtually impossible in the day and age of open access publishing and online repositories of universities, a thorough search has been conducted and 77 published studies, PhD theses and selected MA theses have been analysed with regard to the five dimensions given above.[10] Owing to the diverse perspectives in this emerging research area, not all the empirical investigations mentioned use the term extramural English, they may cover a broader or more specific subject area than EE, or indeed focus on other languages, but in my view they are nonetheless relevant to the present study and can usefully inform further research.

Beginning with the dimension of time, we can immediately see a marked increase in studies in the area of extramural language learning over the past 20 years. Table 2.1 presents all studies included in this review and Figure 2.2 visually displays the number of studies per year. As can be seen, more and more publications began to appear after Benson's (2001) call for research on out-of-class learning mentioned in the introduction to this section and from around 2009 the development picked up pace. Overall, Figure 2.2 shows a steady upward trend and thus continuously growing interest in this new field. This trend also appears to continue in 2019, which promises to be a great year of EE research with a number of highly interesting projects

---

[9] While using a meta-analytic perspective, it is important to highlight that this narrative review does not constitute a meta-analysis in the classic sense as crucial information such as statistical effect sizes (see Plonsky & Oswald 2012) is not taken into account. In fact, studies in this emerging area of enquiry are probably too few and too diverse to allow a statistical meta-analysis as of yet. Thus, rather than drawing inferences from existing research, the goal of this section is to give an overview of what has (not) been done.

[10] MA theses were only included if the object of study is particularly relevant to the present study (i.e. related to vocabulary learning) or if no other research could be found in a given context.

published so far (e.g. De Wilde, Brysbaert & Eyckmans 2019; Hannibal Jensen 2019; Inaba 2019; Jurkovič 2019; Lee 2019a, 2019b; Peters et al. 2019; Puimège & Peters 2019).
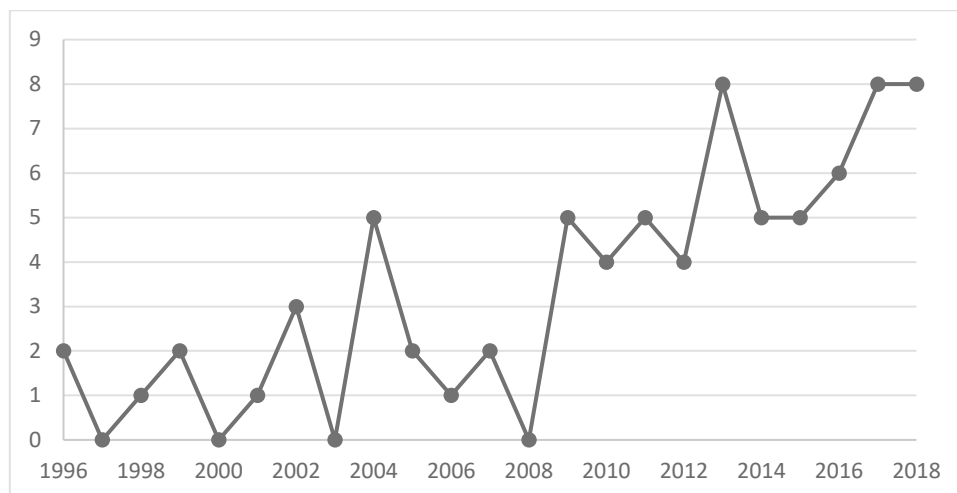
Figure 2.3: Number of empirical studies relating to extramural language learning per year

In terms of location, early research was conducted especially in Asian contexts (e.g. Chan, Spratt & Humphreys 2002; Lamb 2002, 2004a, 2004b; Spratt, Humphreys & Chan 2002; Hyland 2004), while in Europe interest in extramural language learning developed a little later (e.g. Bonnet 2004; Berns, De Bot & Hasebrink 2007; Grau 2009; Lindgren & Muñoz 2013; Peters 2018; Sundqvist 2009a; Sylvén 2004/2010). Research on EE and related concepts is, however, not limited to Asia and Europe, but can by now be found in many different locations worldwide and on all continents including Africa,[11] Australia/Oceania (Pearson 2004), North America (Knight 2007; Lai, Hu & Lyu 2018; Yi 2005) and South America (Cole & Vanderplank 2016). Most studies so far have been conducted in EFL (English as a foreign language) contexts, particularly in Europe and Asia, while fewer studies focus on ESL (English as a second language) settings in the US (Knight 2007; Lai, Hu & Lyu 2018; Yi 2005) or in Asia (Chan, Spratt & Humphreys 2002; Spratt, Humphreys & Chan 2002; Kaur 2015; Lai & Gu 2011; Mukundan, Khojasteh & Pearson 2009). Several European studies have focused on special educational contexts such as Content and Language Integrated Learning (CLIL) (Jakonen 2014; Lancaster 2018; Mirmán Flores & García Jiménez 2018; Olsson & Sylvén 2015; Sylvén 2004/2010, 2006; Verspoor, De Bot & Van Rein 2011). Within Europe another contextual difference that is important concerning research on language learning outside school is the divide between subtitling countries and dubbing countries (see section 1.1). Between 2002 and 2018, 24 empirical studies focusing on learners' EE activities or related concepts were conducted in subtitling countries and areas, such as Finland, Greece, Sweden, the Netherlands or the Flemish part of Belgium, while only 15 studies were found in dubbing countries such as Austria, France, Germany, Spain or Turkey.[12]

---

[11] Dressman, Lee and Sabaoui (2016) present information on students' informal English learning via online media in Morocco in comparison to Korea.

[12] Despite defying any easy assignment to a continent, Turkey is regarded as part of Europe here for reasons of convenience: of the three Turkish studies, one is explicitly located in Istanbul (Ekşi & Aydin 2013) and the second uses the term extramural English (Coşkun & Mutlu 2017), which so far has been used predominantly in northern Europe.

Chapter 2: Extramural English and language learning beyond the classroom

Year	Studies	Number of publications	Year	Studies	Number of publications
1996	Littlewood & Liu (1996) Pickard (1996)	2	1997		0
1998	Yap (1998)	1	1999	Freeman (1999) Suh et al. (1999)	2
2000		0	2001	Pill (2001)	1
2002	Chan, Spratt & Humphreys (2002) Lamb (2002) Spratt, Humphreys & Chan (2002)	3	2003		0
2004	Bonnet (2004) Hyland (2004) Lamb (2004a, 2004b) Pearson (2004)	5	2005	Shen et al. (2005) Yi (2005)	2
2006	Sylvén (2006)	1	2007	Berns, De Bot & Hasebrink (2007) Knight (2007)	2
2008		0	2009	Chusanachoti (2009) Grau (2009) Mukundan, Khojasteh & Pearson (2009) Sundqvist (2009a, 2009b)	5
2010	Inozu, Sahinkarakas & Yumru (2010) Kuppens (2010) Lefever (2010) Sylvén (2004/2010)	4	2011	Anioł (2011) Kalaja et al. (2011) Kuure (2011) Lai & Gu (2011) Verspoor, De Bot & Van Rein (2011)	5
2012	Muñoz (2012) Olsson (2012) Persson & Prins (2012) Sockett & Toffoli (2012)	4	2013	Barbee (2013) Bunting & Lindström (2013) Ekşi & Aydin (2013) Lindgren & Muñoz (2013) Saad, Melor & Embi (2013) Sockett (2013) Sundqvist & Olin-Scheller (2013) Toffoli & Sockett (2013)	8
2014	Bengtsson (2014) Jakonen (2014) Koivistoinen (2014) Liu (2014) Sockett (2014)	5	2015	Kaur (2015) Lai (2015) Lai, Zhu & Gong (2015) Olsson & Sylvén (2015) Trinder (2015)	5
2016	Brevik (2016) Cabot (2016) Cole & Vanderplank (2016) Lai, Yeung & Hu (2016) Nightingale (2016) Olsson (2016)	6	2017	Coşkun & Mutlu (2017) De Wilde & Eyckmans (2017) Ingvarsdóttir & Jóhannsdóttir (2017) Hannibal Jensen (2017) Jóhannsdóttir (2017) Kusyk (2017) Tran (2017) Trinder (2017)	8
2018	Hahn (2018) Lai, Hu & Lyu (2018) Lai & Zheng (2018) Lancaster (2018) Lee & Dressman (2018) Lyrigkou (2018) Mirmán Flores & García Jiménez (2018) Peters (2018)	8			

Table 2.1: Empirical studies on extramural language learning between 1996 and 2018

The backgrounds of the researchers who have recently become interested in language learning outside formal educational contexts are as diverse as the research locations and contexts. The conceptual background of earlier studies frequently lies in learner autonomy (Chan, Spratt & Humphreys 2002; Hyland 2004; Lamb 2004b; Spratt, Humphreys & Chan 2002, see also Benson 2001), but more recent studies also show an interest in autonomous learning (Cole & Vanderplank 2016; Kaur 2015; Liu 2014). Other researchers attempt to investigate success in language learning through the lens of out-of-school activities (Barbee 2013; Brevik 2016; Lai, Zhu & Gong 2015; Lamb 2002) or compare language learning in different contexts (Berns, De Bot & Hasebrink 2007; Bonnet 2004; Tran 2017; Anioł 2011). Studies on the age factor in language learning also take exposure factors more and more into account (Muñoz 2012) and several projects investigate early language learners before the beginning of or within the first years of L2 instruction (De Wilde & Eyckmans 2017; Jóhannsdóttir 2017; Kuppens 2010; Lefever 2010; Lindgren & Muñoz 2013; Persson & Prins 2012). Other researchers studying out-of-class language learning have a background in literacy studies (Kuure 2011; Yi 2005) or focus on researching a particular skill or area of language knowledge (Peters 2018).

Increasingly, however, interest in out-of-class or informal learning appears to be the main motivation for empirical research on its own (e.g. Bunting & Lindström 2013; Hahn 2018; Hannibal Jensen 2017; Lai, Zhu & Gong 2015; Olsson 2012; Pearson 2004; Sundqvist 2009a), or in combination with a focus on learning with technology (Lai & Gu 2011; Lai, Hu & Lyu 2018; Lai & Zheng 2018; Liu 2014) or in online contexts (Kusyk 2017; Lee & Dressman 2018; Sockett 2013, 2014; Sockett & Toffoli 2012; Trinder 2015, 2017). Individual studies are beginning to look at learner variables and context factors mediating extramural activities, such as family background (Lindgren & Muñoz 2013; Mirmán Flores & García Jiménez 2018), agency (Kalaja et al. 2011; Lyrigkou 2018) or motivation (Barbee 2013; Lamb 2004a). Similarly, the relationship between language learning inside and outside school (Bunting & Lindström 2013; Jakonen 2014; Lai 2015; Saad, Melor & Embi 2013) as well as at learners' and teachers' perceptions of and views on learners' out-of-school language activities (Grau 2009; Lai, Yeung & Hu 2016; Sundqvist & Olin-Scheller 2013; Toffoli & Sockett 2013) are receiving more attention.

Depending on the different backgrounds and research interests, many different concepts are used to describe the object of investigation in these studies. As can be seen from Table 2.2, variations of out-of-class or out-of-school learning are clearly in the lead with 23 studies using such terms, followed by 13 studies using the concept of extramural English, seven studies describing their object of study as online informal learning of English (OILE) and six studies using exposure as their central term. Other terms and concepts that were used in more than one study include informal learning, extracurricular exposure or input, and incidental language learning. In addition, several more recent concepts, such as IDLE or FASILS, have only be used in one study before 2019.

Concept	Studies	N of studies
out-of-class/out-of-school language learning	Chusanachoti (2009) Cole & Vanderplank (2016) Ekşi & Aydin (2013) Grau (2009); Hyland (2004) Inozu, Sahinkarakas & Yumru (2010) Lai (2015); Lai & Gu (2011) Lai, Hu & Lyu (2018) Lai, Yeung & Hu (2016) Lai & Zheng (2018); Liu (2014) Mukundan, Khojasteh & Pearson (2009) Nightingale (2016); Pearson (2004) Saad, Melor & Embi (2013) Shen et al. (2005)	17
out-of-class English exposure/ out-of-school language input	Knight (2007) Verspoor, De Bot & Van Rein (2011)	2
out-of-school practices/ out-of-school literacy practices	Bunting & Lindström (2013) Yi (2005)	2
outside class activities	Chan, Spratt & Humphreys (2002) Spratt, Humphreys & Chan (2002)	2
extramural English/ extramural activities/ extramural exposure	Bengtsson (2014) Coşkun & Mutlu (2017) Hahn (2018) Ingvarsdóttir & Jóhannsdóttir (2017) Hannibal Jensen (2017); Lancaster (2018) Olsson (2012, 2016) Olsson & Sylvén (2015) Sundqvist (2009a, 2009b) Sundqvist & Olin-Scheller (2013) Sylvén (2006)	13
OILE	Kusyk (2017) Sockett & Toffoli (2012) Sockett (2013, 2014) Toffoli & Sockett (2013) Trinder (2015, 2017)	7
intensive exposure/ (out-of-school) exposure/ exposure to target language outside the classroom	Jóhannsdóttir (2017) Lindgren & Muñoz (2013) Mirmán Flores & García Jiménez (2018) Muñoz (2012) Persson & Prins (2012) Peters (2018)	6
informal learning/ learning through informal contact	Jakonen (2014); Lyrigkou (2018) Tran (2017)	3
extracurricular exposure to English/ extracurricular L2 input	Anioł (2011) Barbee (2013)	2
incidental language acquisition/ incidental language learning	De Wilde & Eyckmans (2017) Kuppens (2010)	2
everyday learning	Koivistoinen (2014)	1
FASILs	Cole & Vanderplank (2016)	1
IDLE	Lee & Dressman (2018)	1
independent language learning	Lamb (2004b)	1
language learning activities outside the classroom	Lamb (2002)	1
language learning beyond the classroom	Lai & Zheng (2018)	1
learning ecologies/PLEs	Cabot (2016)	1
technology-mediated language learning	Kuure (2011)	1

Table 2.2: Concepts used to describe the object of study in studies on informal out-of-school language learning

In terms of the target language being learned, the vast majority of studies focus on English only. Three projects compare extramural learning of English with other languages such as Bahasa Malaysia (Mukundan, Khojasteh & Pearson 2009), Catalan and Spanish (Nightingale 2016), Swedish (Kalaja et al. 2011), or French or Spanish for learners living in the UK (Lindgren & Muñoz 2013). Only four studies do not include English: Bengtsson (2014) investigates extramural activities of learners of Japanese in Sweden, and Lai (2015), Lai, Hu and Lyu (2018) and Lai and Zheng (2018) incorporate diverse target languages including German, Korean and Spanish.

In contrast, the content focus of the empirical studies varies more widely. Several projects are concerned with mapping a group of learners' out-of-class activities and/or resources (Cabot 2016; Hyland 2004; Chusanachoti 2009; Ekşi & Aydin 2013; Inozu, Sahinkarakas & Yumru 2010; Lai, Zhu & Gong 2015; Mukundan, Khojasteh & Pearson 2009; Pearson 2004; Saad, Melor & Embi 2013; Shen et al. 2005) or closely describe the practices of one learner (Koivistoinen 2014; Kuure 2011) or a small group of learners (Brevik 2016; Sockett & Toffoli 2012). In addition, a number of studies attempt to empirically relate extramural practices to language learning, for instance, to overall language proficiency (Bengtsson 2014; Berns, De Bot & Hasebrink 2007; Bonnet 2004; Cole & Vanderplank 2016; De Wilde & Eyckmans 2017; Verspoor, De Bot & Van Rein 2011) or to language skills. Some of these focus on all four language skills (Coşkun & Mutlu 2017), but the majority of the skills-focused studies concentrate on one or two skills, such as listening (Kuppens 2010; Lefever 2010), reading (Lefever 2010; Verspoor, De Bot & Van Rein 2011; Yi 2005), writing (Olsson 2012, 2016; Verspoor, De Bot & Van Rein 2011; Yi 2005) or speaking (Kuppens 2010; Lee & Dressman 2018; Lefever 2010; Lyrigkou 2018; Sundqvist 2009a, 2009b). While it may appear that there is an emphasis on investigating productive skills, this would not be the case if studies focusing on one particular activity such as reading or viewing audiovisual media (see section 3.3.2) were taken into account as well. While studies linking extramural practices to specific language skills are still relatively scarce, there is by now a substantial body of research on out-of-school activities and vocabulary acquisition (Berns, De Bot & Hasebrink 2007; De Wilde & Eyckmans 2017; Hahn 2018; Hannibal Jensen 2017; Jóhannsdóttir 2017; Kaur 2015; Lee & Dressman 2018; Olsson 2016; Olsson & Sylvén 2015; Persson & Prins 2012; Peters 2018; Sundqvist 2009a, 2009b; Verspoor, De Bot & Van Rein 2011), which will be described in greater detail in section 3.3.1. In contrast, only one study takes grammatical knowledge into account (Persson & Prins 2012). In addition to skills and areas of language knowledge, other studies focus on contextual factors, the link between EE and formal instructed contexts, or special educational contexts such as CLIL, as mentioned above.

Concerning the research designs used to investigate these diverse foci, many studies use mixed method designs combining quantitative and qualitative perspectives. Overall, there appears to be greater emphasis on the quantitative perspective with ten studies using a questionnaire as the only (e.g. Ingvarsdóttir & Jóhannsdóttir 2017; Mukundan, Khojasteh & Pearson 2009; Shen et al. 2005) or one of several instruments, for instance in combination with one-on-one or group interviews (e.g. Barbee 2013; Chan, Spratt & Humphreys 2002; Inozu, Sahinkarakas & Yumru

2010; Lai & Gu 2011). In contrast, few studies follow an in-depth qualitative approach concentrating on one individual (Koivistoinen 2014; Kuure 2011) or a small group of learners (Brevik 2016; Kusyk 2017; Sockett & Toffoli 2012). Studies attempting to relate extramural practices to (aspects of) language development commonly include various types of tests and/or self-assessment tools (e.g. Berns, De Bot & Hasebrink 2007; Kuppens 2010; Lindgren & Muñoz 2013; Peters 2018). Many studies also make use of some form of (structured) diary or activity log to record out-of-school practices in greater detail (e.g. Anioł 2011; Olsson 2012; Sundqvist 2009a; Yi 2005). In addition, a wealth of other data collection methods has been used including, but not limited to, reflective journals, observations, audio and video recordings, blog entries and other online data.

The majority of studies present cross-sectional snapshots of learners' out-of-school practices, but a few projects provide more longitudinal perspectives (e.g. Chusanachoti 2009; Kuure 2011; Sundqvist 2009a; Verspoor, De Bot & Van Rein 2011; Yi 2005). In terms of participant groups, there is an emphasis on secondary and tertiary contexts, with fewer studies focusing on children attending primary school (De Wilde & Eyckmans 2017; Hannibal Jensen 2017; Jóhannsdóttir 2017; Kuppens 2010; Lefever 2010; Lindgren & Muñoz 2013; Persson & Prins 2012) or on adult learners (Hyland 2004; Knight 2007). Sample sizes vary strongly according to the studies' aims and methods used; for instance, among the 77 projects included in this review there are 11 qualitatively-oriented studies with fewer than ten participants, but at the same time there are 24 quantitatively-oriented studies with more than 200 participants and three large-scale multi-site projects involving samples from several European countries (Berns, De Bot & Hasebrink 2007; Bonnet 2004; Lindgren & Muñoz 2013).

To sum up, this review of 77 empirical studies belonging to the growing body of research on language learning beyond the classroom suggests that interest in the area of enquiry as well as the number of studies is growing steadily and that by now research on this topic is carried out in many, highly diverse contexts around the world. English clearly is the predominant target language and a much larger number of studies is carried out in EFL than in ESL contexts. Researchers from various backgrounds, such as learner autonomy, age-related research and early language learning, as well as learning with technology and in digital contexts have become interested in this field, resulting in many different terms and theoretical concepts being used to describe the object of investigation. While some empirical studies primarily aim to map language-related practices outside formal educational contexts, others connect extramural practices and language development by focusing on overall language proficiency, language skills or vocabulary knowledge. A consequence of these diverse backgrounds, concepts and foci is the use of a variety of different research designs, although cross-sectional quantitative perspectives currently represent the majority of empirical studies.

What becomes clear from this overview along five meta-analytic dimensions is that there is still much work to be done. First and foremost, there is a lack of studies targeting languages other than English. Second, there is a predominance of research in Europe and Asia with other

contexts in danger of being neglected. Third, some theoretical consolidation and further development of models such as Benson's (2011) framework would help to increase comparability between studies and reduce terminological confusion, since at the moment many researchers seem to simply invent their own term or give relatively vague descriptions of what it is they are studying. Fourth, further, and in some cases perhaps more rigorous, empirical research is needed to relate extramural activities to language learning and to allow conclusions on the benefits for language knowledge or skills to be drawn. At the same time, more in-depth qualitative studies are needed to better understand learners' perspectives and to gain a clearer idea whether, to what extent and how out-of-school language learning can be connected to in-school learning and teaching (see also Chapter 9).

2.4   Summary

This chapter presented the current state of theories and research on informal learning outside educational institutions with a focus on extramural English and language learning beyond the classroom. In the emerging research field, many competing concepts and terms are currently used, but these two conceptualizations are the most fitting for the present study. Language learning beyond the classroom is a wider perspective that includes many different kinds of learning situations, instructed or non-instructed, self-directed or other-directed, formal or informal. Nonetheless, this approach is of great value as Benson's (2011) model is the first attempt to provide a coherent theoretical framework for the developing research area, which can be used to characterize different contexts. In contrast, extramural English presents a narrower perspective on the one hand because it focuses on informal, learner-driven leisure activities involving English, but on the other hand, it can also be regarded as a broader conceptualization than LBC because it is interested in language use rather than language learning. EE includes all informal activities in which learners use English and in which learning is usually not the primary focus. Hence, sometimes an EE activity may lead to language development and sometimes the same activity may not. In short, this study is interested in extramural English defined as English-language activities which take place outside the walls of educational institutions in Austria during learners' leisure time and which are learner-driven and typically voluntary and informal.

Next, an overview of existing EE research along the five meta-analytic dimensions of time, location, conceptual background, content focus and research design showed that interest in informal language learning has rapidly increased over the last decade. By now, research has been conducted in very different contexts all over the world and similarly, researchers from various background have contributed to the existing body of literature. This also results in a wide range of different terms and theoretical concepts being used to describe the object of investigation, which is, however, also the case in the wider field of research on informal learning (see section 2.1). While there is also great variety in terms of content focus and methodology, the trends are clearer in this respect with most studies mapping out-of-school language environments and/or investigating links between EE activities and language development. In

addition, there is a strong focus on English as the target language with few studies investigating other languages. In terms of research design, cross-sectional quantitative perspectives are currently in the lead, but again there is great diversity overall. What became clear from this review of previous studies is that there is much work to be done in relation to linguistic diversity, theoretical development, and rigorous and complementary methodological approaches.

3 Vocabulary learning in extramural contexts

While research on language learning in the twentieth century largely focused on grammar and syntax, interest in lexical issues began to grow in the latter half of the century (see also Barcroft, Sunderman & Schmitt 2011). As Meara's (2012, 2014, 2015b, 2016, 2017) bibliographical analyses of early vocabulary research show, "a distinctive L2 vocabulary research program is beginning to emerge" (Meara 2017: 151) in the 1980s and the publication of several monographs in the 1990s, most notably Paul Nation's volume on *Teaching and Learning Vocabulary,* marked the beginning of a prolific period of lexical research (Laufer & Nation 2012), which turned the field into the thriving research area it is today .

As shown in Chapter 2, within the larger field of lexical research there is by now a small body of research on the specific topic of this thesis: vocabulary learning in extramural contexts. The main aim of this chapter is to provide a more detailed review of vocabulary acquisition from engagement with extramural English and vocabulary uptake from specific activities such as reading or listening, but first foundational notions in relation to L2 vocabulary knowledge and development as well as issues related to vocabulary testing are introduced in order to contextualize the studies presented in section 3.3. For this reason, section 3.1 provides information about different constructs of vocabulary knowledge, about research on L2 lexical development and its learning goals, processes, and outcomes in relation to vocabulary size, as well as about factors known to influence vocabulary learning from input. In section 3.2, methods and issues of vocabulary measurement are discussed because testing methods have a crucial influence on research outcomes and their interpretation. Finally, section 3.3 presents a detailed review of vocabulary learning from EE in general (section 3.3.1) and from individual activities such as reading, listening, viewing and gaming (section 3.3.2).

3.1 Vocabulary development

This section aims to provide a brief overview of the foundations of vocabulary research and critical issues (section 3.1.1), before synthesizing the current state of knowledge on lexical development in the L2 with a focus on vocabulary size (section 3.1.2) and discussing the influence of various factors on L2 vocabulary acquisition (section 3.1.3). While it is impossible to provide a detailed account of all basic concepts within the space available, awareness of fundamental issues and positions in the research field is necessary to contextualize the findings reported later in this chapter.

3.1.1 What it means to know a word

Although the succinct subtitle above may seemingly suggest so, neither the notion of a 'word' nor what it means to 'know' a word is unproblematic; in fact, as surprising as it may seem there is no generally agreed upon definition of what counts as a word in the field of vocabulary research (see also Gardner 2007). Several ways of counting lexical items such as tokens, types, lemmas (Francis & Kučera 1982), flemmas (Pinchbeck 2014 as cited in McLean 2018) and word

families (Bauer & Nation 1993) have been proposed and made use of in different areas of study. However, the conceptualization and counting of lexical items is a central issue in much vocabulary research because it influences vocabulary measurement, empirical results and their interpretation. Taking the base words 'develop' (Bauer & Nation 1993: 254) and 'fish' as examples, Table 3.1 demonstrates that the different units of counting differ greatly in their scope and thus the decision on which to use has an immense impact on the results of lexical research, most notably on frequency lists established on the basis of large corpora.

Unit of counting	develop	fish
Lemma (Francis & Kučera 1982)	develop (verb), develops, developed, developing	fish (verb), fishes, fished, fishing
Flemma (McLean 2018)	develop (verb), develops, developed, developing	fish (verb), fishes, fished, fishing, fish (noun),
Word family (WF 6) (Bauer & Nation 1993)	develop (verb), develops, developed, developing, developer(s), development(s), undeveloped, developmental, developmentally, redevelop, semideveloped	fish (verb), fishes, fished, fishing, fish (noun), fishy

Table 3.1: Examples of different units of counting

Traditionally, word families have been used for frequency lists (e.g. Nation 2006, 2012a and well-known tests (e.g. VLT, LVLT, or VST, see section 3.2.3) and they have dominated some areas such as research on comprehension and coverage (see section 3.1.2). Recently, this practice has been called into question because evidence for the adoption of word families largely comes from L1 studies (Bertram, Laine & Virkkala 2000; Nagy et al. 1989; Wysocki & Jenkins 1987), while several studies have shown that it is not the best counting unit for research concerned with L2 vocabulary learning.[13] L2 learners are not always able to reliably connect a known base word to its derivative forms (Kremmel & Schmitt 2016), produce derivative forms of a given base word (Schmitt & Zimmermann 2002), or correctly translate both the base word and a derivation (Ward & Chuenjundaeng 2009). Clearly, this is problematic and therefore researchers such as Schmitt (2010) or Kremmel (2016) argue for the adoption of the lemma as counting unit because it is the simpler and more transparent option that allows the direct comparison of receptive and productive vocabulary size[14] and helps to prevent underestimation of vocabulary sizes needed to accomplish communicative tasks.[15] Another possibility advocated by McLean (2018) is the flemma, which is similar to the lemma in that it includes all inflectional forms, but it does not differentiate between parts of speech. McLean (2018) regards flemmas as more ecologically valid because word lists based on computerized corpus analysis commonly do not distinguish between different parts of speech. It is certainly worth noting that the units of counting used in

---

[13] It is important to point out that word families have traditionally been operationalized at level 6 (McLean 2018), the last and most inclusive level of Bauer and Nation's (1993) six stages of building word families, although the authors stress the evolving nature and families in the original article. This operationalization of word families at the most inclusive level is particularly problematic in relation to research on L2 learners.

[14] The word family is generally advocated only in relation to receptive vocabulary knowledge, for instance Bauer and Nation (1993) explicitly refer to reading in their seminal article.

[15] Schmitt (2010) argues that there is a danger that consumers of vocabulary research may simply equate word families with words, which would lead to gross underestimation of the amount of vocabulary needed for comprehension as found in coverage research, which typically has used word families as units of counting (see section 3.1.2).

lexical research are influenced by computational possibilities because research findings, such as coverage figures, as well as research instruments, such as vocabulary tests, are based on frequency lists drawn from large corpora with the help of automated counts. Machine-based frequency counts are useful, but subject to certain restrictions because they cannot automatically distinguish between different parts of speech or multiple meanings without elaborate preparation of corpora through tagging.

Multiple meanings and formulaic language present additional challenges regarding the definition of a word (Gardner 2007). Although the importance of formulaic language or multiword items is gradually becoming clearer in research (for a synthesis see Wray 2012, 2013) and is highlighted by prominent vocabulary researchers (e.g. Schmitt 2010: 8–12), the impact on vocabulary teaching and testing is still limited. Because formulaic language itself as well as research on formulaic language is a multi-faceted and varied phenomenon, it is difficult to integrate formulaic language or multiword items into traditional approaches to vocabulary research, which have commonly focused on individual lexical items.[16] Due to the emerging nature of research on formulaic language, the computational limitations mentioned above, and a lack of consensus among researchers, frequency lists and vocabulary size tests based on these lists do not currently take multiword items into account (Laufer & Nation 2012: 164). However, interest in L2 learners' acquisition of multiword items is growing (Wray 2012: 236) and there are first attempts to establish lists of multiword items for teaching purposes and as a basis for integrating them into vocabulary tests (e.g. Garnier & Schmitt 2015; Martinez & Schmitt 2012). In sum, at present there is no generally accepted definition of what counts as a word in vocabulary research. As Wray (2012: 248) puts it "[w]ords are more elusive than they ought to be" and it is thus of great importance among both researchers and practitioners to clearly acknowledge what exactly it is they mean by a 'word'.

In addition, 'knowing' a word is as problematic a notion as a word itself, and Milton and Fitzpatrick (2014: 1) seem to echo Wray's quote above in stating that "[k]nowing a word is an elusive concept and we are still unable to capture, in a simple description, everything that knowing a word might involve". In the literature there are three approaches that attempt to shed light on the concept of word knowledge: the component approach, which is most famously represented by Nation's (2001) taxonomy, the developmental approach, with the Vocabulary Knowledge Scale (VKS) by Paribakht and Wesche (1993, 1997) as the most widely known example, and the metaphorical approach, which is used, for instance, in the lexical space metaphor proposed by Daller, Milton and Treffers-Daller (2007) and in conceptualizations of lexical networks (e.g. Aitchison 2003; Meara 1996; Meara & Wolter 2004).

---

[16] Formulaic language is a varied phenomenon, examples of multiword items or formulaic sequences include idioms, proverbs, phrasal verbs, collocations and many other types. The diversity of this field is also apparent in Wray's (2002: 9) definition of a formulaic sequence, which is another of the many terms used to label lexical phrases consisting of more than one element, "as a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar."

In comparison to the other two approaches the component approach is, however, the most well-developed (Milton & Fitzpatrick 2014) and also the most widely-used, which is why it constitutes the focus of this section. As exemplified by Nation's (2001) taxonomy, component approaches list different aspects of word knowledge, often contrasting them with each other. The earliest example of such a contrast is the differentiation between receptive and productive mastery, which appears to have been introduced by Palmer in 1921 (Milton & Fitzpatrick 2014: 2). More detailed lists of word knowledge aspects begin to appear in the 1940s (Cronbach 1942) and are elaborated on in the 1970s (Richards 1976), culminating in the taxonomy proposed by Nation (2001), which is the most comprehensive version to date. It includes 18 different aspects of word knowledge related to the three main categories of form, meaning and use and thus illustrates that knowing a word is far from simple. The taxonomy also highlights that word knowledge is a complex construct, but it is worth noting that recent research by González-Fernández and Schmitt (2019) indicates that explicit vocabulary knowledge is a unidimensional construct; hence, González-Fernández (2019) suggests that the different knowledge components listed by Nation should be regarded as levels of one construct rather than as separate entities.

A further attempt to specify word knowledge, which can be seen as part of the component approach, are the so-called dimensions of knowledge first introduced by Anderson and Freebody (1981). Their contrast between the *breadth* or size of a person's vocabulary and the *depth* of their knowledge, i.e. what they know about each lexical item, has also had great impact on both theoretical discussions and research design. The depth dimension, however, has been conceptualized and operationalized in different ways by different scholars, which led Read (2004) to distinguish between three different types of depth of knowledge: precision of meaning, comprehensive word knowledge and network knowledge. The dimensional approach to vocabulary knowledge has also been extended by a third dimension: Daller, Milton and Treffers-Daller (2007: 8) add a fluency dimension that "is intended to define how readily and automatically a learner is able to use the words they know and the information they have on the use of these words" to form the *lexical space*. A further conceptualization of the dimensions of vocabulary knowledge has been proposed by Meara (1996: 48), who suggests that instead of focusing on individual words, we should concentrate on the "degree of connectivity" in the lexicon and replace the dimension of depth with *organization,* which is characterized by the links among words in the lexicon. As mentioned above, both of these extensions are related to metaphorical approaches to vocabulary knowledge and are also concerned with the fluent retrieval of word knowledge, which has not been a focus in the word list approach.

Although the word list approach and the dimensions approach conceptualize word knowledge as consisting of different components, be it more specific knowledge aspects or larger dimensions, it is unclear how they relate to each other (Milton & Fitzpatrick 2014). While the breadth or size dimension seems to include knowledge of word form and potentially the form-meaning link, knowledge of word parts, which Nation (2001) includes as part of knowledge of form, would probably rather be included in the depth dimension along with further, more fine-

grained aspects of knowledge relating to meaning and use.[17] The basic problem underlying this issue is that there is no agreed upon definition what exactly constitutes the most basic form of knowledge; or in other words at which point a word is sufficiently known to form part of size and which aspects of knowledge can then be added onto this basic knowledge as part of the depth dimension. What does, however, become clear from this discussion is that simply referring to lexical items as known or unknown is an unnecessary imprecision and "wholly inadequate for describing vocabulary knowledge" (Schmitt 2010: 22). Instead, an effort should be made to state precisely which aspect of word knowledge has been established for which type of lexical item at the time of measurement since, as we shall see in the next section, vocabulary acquisition is not linear and subject to phenomena of attrition like any other type of learning.

This section addressed foundational concepts of vocabulary research and showed that several unresolved issues are hidden behind the deceptively simple phrase 'knowing a word'. Neither the construct of a 'word' nor that of 'knowing' one is unproblematic, and this brief overview highlighted that there are many approaches to defining either in lexical research. In relation to the construct of a word, different units of counting have been proposed with many researchers currently seeming to lean more towards smaller units, such as the lemma or the flemma. At the same time, larger lexical units consisting of more than one word form, such as multiword items and formulaic language, pose a challenge in terms of counting that has not been resolved as of yet. Although the importance of formulaic language is becoming increasingly clear, it has not yet been included in conventional frequency lists and standardized tests of vocabulary knowledge.

Similar to units of counting there are many conceptualizations of word knowledge with the component approach emerging as the most well-researched and thus most influential approach to word knowledge. However, as Milton and Fitzpatrick (2014: 1) point out:

> The development of our understanding over the last 100 years or so has entailed breaking down the idea of word knowledge into progressively smaller and smaller areas. This has made the analysis of the subject progressively more complex, but not necessarily clearer.

While the fragmentation of word knowledge allows much more fine-grained analyses, it also poses practical problems for research and test design (e.g. the need for multiple measures and resulting testing effects). Since vocabulary measurement in this project focuses on size rather than depth, such issues relating to the testing of multiple word knowledge components will not be discussed in detail, but an overview of problems in vocabulary size testing is given in section 3.2.

---

[17] The practices in vocabulary testing, where size tests have traditionally tested recognition of word form and/or recognition of the form-meaning link, appear to support the conclusion that knowledge of forms and the basic form-meaning link are part of vocabulary breadth; see also Daller, Milton and Treffers-Daller (2007) for their conceptualization of the relation between dimensions and word knowledge aspects as part of the lexical space metaphor.

3.1.2 Lexical development in the L2: processes, goals and outcomes

Vocabulary learning is both a quantitative and a qualitative challenge for L2 learners (Laufer & Nation 2012: 163): first, there are thousands of words that are waiting to be learned and second, many different aspects of knowledge should be mastered for each of them (see section 3.1.1). It is therefore clear that vocabulary learning is an ongoing, lifelong task – especially since the lexicon of any language is constantly subject to change (see also Dóczi & Kormos 2016: 1).[18] In the following, this section provides a brief introduction to the fundamentals of L2 vocabulary development and synthesizes literature on vocabulary learning processes, learning goals, and learning outcomes.

As pointed out above, using a language effectively requires (some) knowledge of many different words and multiword items, which cannot be acquired all at once. Similarly, developing knowledge of a specific word or phrase requires more than one, and often numerous, encounters. Vocabulary learning is therefore incremental "both in terms of acquiring an adequate vocabulary size, and in terms of mastering individual lexical items" (Schmitt 2010: 19, see also Schmitt 2000). In addition, the acquisition of each word knowledge aspect itself can be conceptualized as developing on a cline from partial to precise knowledge (Henriksen 1999). Thus, lexical development is incremental in at least three related ways, and since many encounters are needed to develop the lexicon and word knowledge, vocabulary learning is strongly related to frequency (Ellis 2002). It is well established that high frequency vocabulary is learned earlier than lower frequency vocabulary (Read 1988; Schmitt, Schmitt & Clapham 2001) and lexical items with a higher frequency are also more useful for language comprehension and use (Nation 2013). There is a well-established relationship between word frequency and lexical coverage known as Zipf's law (Milton 2009), which shows that a very small proportion of highly frequent words provide large amounts of coverage in any given text. On the basis of this relation, researchers differentiate between different frequency levels; for example, Nation (2013) distinguishes between high-frequency (2,000 word families, 1K-2K), mid-frequency (7,000 word families, 3K-9K) and low-frequency words (>9K) based on his BNC/COCA lists (Nation 2012a).[19, 20]

In addition, researchers have argued that some aspects of lexical knowledge are acquired earlier than others and most agree that recognizing a word form and connecting meaning to it "is the first and most essential lexical aspect which must be acquired" (Schmitt 2010: 15, see also Ma 2009; Nation 2007). Research has shown that receptive knowledge generally precedes productive knowledge (Dóczi & Kormos 2016: 69–81; González-Fernández & Schmitt 2019). In addition, word form operationalized as spelling appears to be the easiest word knowledge aspect to acquire

---

[18] Brysbaert et al. (2016) provide evidence from a massive online experiment, which shows that vocabulary size increases with age and that vocabulary acquisition is an ongoing, lifelong task also for L1 users.

[19] Frequency lists have traditionally been split into 1,000-item bands (for a critique see Kremmel 2016), which are often abbreviated using the letter K; thus, 2K refers to the 2,000 most frequent words of a given frequency list.

[20] Until quite recently, a distinction was commonly only made between high- and low- frequency vocabulary; however, Schmitt and Schmitt (2014) and Nation (2013) present a convincing case for introducing a third level of mid-frequency vocabulary.

(Schmitt 1998) and knowledge of the form-meaning link is acquired before aspects such as derivatives and polysemy (Dóczi & Kormos 2016; González-Fernández & Schmitt 2019). While the nature of results in such research quite obviously depends on the exact operationalization of the word knowledge aspects, these studies provide initial evidence that there may be an order of acquisition for lexical knowledge and show that some aspects of word knowledge are typically acquired before others.

In relation to receptive and productive knowledge, similar results have also been found in research on vocabulary size. Using the CATSS (see section 3.2.3), which measures different levels of the form-meaning link, Laufer and Goldstein (2004) showed that there was a stable hierarchy of difficulty with meaning recognition and form recognition being easier than meaning recall and form recall.[21] Using the same data, Laufer (2005: 232) expresses the difference between productive size, operationalized as form recall, and receptive size, measured through meaning recall, as a ratio. She shows that the productive/receptive ratio varies between 35% at the 2K frequency level and 16% at the 5K level. This indicates two things: first, even at the level of high-frequency words (2K) the mean productive vocabulary size (M = 6.51) was only 35% of the receptive vocabulary size (M = 18.61) and second, with decreasing word frequency the productive/receptive ratio decreases, meaning that even fewer words are known productively at lower frequency levels. Based on this study and further research with similar results (Laufer & Paribakht 1998; Tschirner 2004; Waring 1997), we can be reasonably sure that learners' receptive vocabulary size is larger than their productive one.

However, learners are unlikely to have stable vocabulary sizes because "[v]ocabulary learning is not a tidy linear affair, with only incremental advancement and no backsliding" (Barcroft, Sunderman & Schmitt 2011: 578). Like any other type of knowledge, lexical knowledge is subject to attrition, and research suggests that vocabulary knowledge may be affected even more by attrition phenomena than other areas of linguistic knowledge (Schmitt 2010: 23). In addition, studies have found that productive knowledge is more prone to attrition than receptive knowledge (Bardovi-Harlig & Stringer 2010; Weltens & Grendel 1993) and that influencing factors can have different effects on processes of vocabulary learning and decay (Barclay 2017, 2018).

Next, we turn to the processes of vocabulary learning, focusing on the distinction between incidental and intentional learning, which is of particular interest in relation to learning in out-of-school contexts. This distinction has been widely debated in the field of lexical research and incidental learning has received much attention as it seems highly unlikely that the thousands of words needed to use an L2 effectively can be acquired through intentional study alone (e.g. Gass 1999, see below).

---

[21] For an explanation of the terms form/meaning recall and recognition, please see section 3.2.1. The hierarchy found by Laufer and Goldstein (2004) was meaning recognition < form recognition < meaning recall < form recall, making form recall the most difficult aspect of knowledge with regard to the form-meaning link. In their study on incidental learning from reading, Eckerth and Tavakoli (2012) found a slightly different order of difficulty with form recognition < meaning recognition< meaning recall < form recall, but the general trend of receptive knowledge being more easily acquired than productive knowledge remains the same.

While originally the distinction between intentional and incidental learning referred to a methodological criterion, the term incidental learning has subsequently been used as a theoretical construct in SLA to refer to learning "without the conscious intention to commit the element to memory" (Hulstijn 2013: 1) and has often been conceptualized as a by-product of language use (e.g. Huckin & Coady 1999; Wode 1999).[22] The distinction between incidental and intentional learning is frequently related to another pair of contrasting learning processes relating to Schmidt's (1994) concept of consciousness as awareness: implicit and explicit learning. Implicit learning is defined as "learning without awareness of what is being learned" (DeKeyser 2003: 314, see also Rebuschat & Williams 2013), while explicit learning refers to a "conscious operation where the individual makes and tests hypotheses in a search for structure" (Ellis 1994b: 1, see also Muñoz 2013). While some authors see a large overlap between the pairs of incidental/intentional and implicit/explicit learning (e.g. Hulstijn 2003), Rieder (2003) and more recently Ender (2016) have argued that incidental vocabulary learning can involve both implicit and explicit processes, depending on the awareness of learning on the part of the learner.[23]

With regard to vocabulary, incidental learning has mostly been researched in relation to reading (e.g. Elgort et al. 2018; Elgort & Warren 2014; Malone 2018; Mohamed 2018; Pellicer-Sánchez 2016; Pellicer-Sánchez & Schmitt 2010; Pigada & Schmitt 2006; Schmitt, Jiang & Grabe 2011; Waring & Takaki 2003; Webb & Chang 2015), while studies on incidental learning from listening are much rarer (e.g. Brown, Waring & Donkaewbua 2008; Van Zeeland & Schmitt 2013a; Vidal 2003, 2011). Generalizing very broadly, these studies have found that incidental learning does occur, but gains are typically rather small (see also section 3.3.2 for a more detailed review of studies relevant to this project). However, as argued by Nation (2013: 356): "Small gains can become large gains if learners do large quantities of reading [or listening]." Furthermore, contextual word learning from reading seems to be more beneficial for some types of word knowledge than others, i.e. it has stronger effects on word form than on word meaning (Schmitt 2010: 29–30). Despite the value of incidental learning, it is becoming increasingly clear that "naturalistic usage-based learning [on its own] is insufficient to acquire second-language (L2) vocabulary [...] and needs to be supplemented by deliberate form-focused learning" (Elgort 2011: 368). In this respect, research has shown that deliberate or intentional learning is particularly effective for the relatively small amount of high frequency words or for specialized vocabulary items (Elgort 2011: 400; Nation 2013: 22–28).

---

[22] First used in psychological experiments, the term intentional learning indicated that participants were explicitly instructed to study or memorize target structures and informed they would be tested afterwards, while in an incidental learning condition participants were not made aware of a subsequent test (Hulstijn 2003).

[23] Types of learning processes, such as incidental and intentional learning or implicit and explicit learning, are linked to the notions of consciousness, awareness and attention. In his seminal article, Schmidt (1994) distinguished between four different senses of the word consciousness: consciousness as intentionality, as attention, as awareness, and as control. These four senses can be related to different types of learning processes. The notion of consciousness as intentionality refers to the distinction between incidental and intentional learning; thus, to learning with or without the intent to learn, whereas explicit and implicit learning related to consciousness as awareness.

In contrast to research on incidental and intentional vocabulary learning, which has a relatively long tradition in the field, the implicit/explicit distinction has only recently received increased interest. For the most part, vocabulary tests measure explicit knowledge of form, meaning or use (Elgort 2011) and research into implicit lexical knowledge has only recently begun to emerge.[24] Sonbul and Schmitt (2013: 125) argue that because vocabulary acquisition has commonly been regarded as a form of item learning rather than rule learning, it has been assumed that "vocabulary knowledge is declarative in nature and can never be implicit". They further report that Ellis (1994a) was one of the first to distinguish between word knowledge aspects which can best be learned explicitly, such as meaning and semantic relations, and other features, such as form and aspects of use, which are best learned implicitly. Elgort (2011) was one of the first to empirically investigate the role of implicit and explicit knowledge in relation to vocabulary and showed that deliberate learning of nonwords from word cards led to both declarative and procedural knowledge, as measured through offline tests and online priming tasks. Other studies have begun to investigate implicit and explicit knowledge of collocations (Sonbul & Schmitt 2013) or effects of contextualized learning on the development of implicit and explicit knowledge (Choi, Kim & Ryu 2014; Elgort & Warren 2014). While the measurement of implicit knowledge is still far from common in lexical research, interest in this area, which may have wider implications for the conceptualizations of knowledge in the field (Godfroid 2019), is growing rapidly.

This brief account has introduced some fundamental aspects of L2 vocabulary development, such as incrementality, the relation to frequency, first evidence for an order of acquisition in relation to word knowledge aspects, the evident presence of attrition, the roles of intentional and incidental learning, and the development of explicit and implicit knowledge. Yet, it also shows that there are large gaps in our knowledge and that we have no overall theory of vocabulary development. This gap is prominently lamented by Schmitt (2010: 26):

> While we are gaining an increasing understanding of the development of some isolated aspects of vocabulary, the overall acquisition system is far too complex and variable for us to comprehend it in its entirety, and so still eludes description.

Schmitt (2019: 261) also positions the development of "a practical model of vocabulary acquisition" as the first and most prominent of nine research tasks in a recent research agenda for the whole field. He suggests extending the study by González-Fernández and Schmitt (2019) with further word knowledge components as well as longitudinal studies as a way forward to address this issue and arrive at a comprehensive theory of vocabulary learning.

In addition to the processes of vocabulary development, the goals and outcomes of lexical learning also need to be addressed here. A claim that has been encountered several times so far is that knowledge of many lexical items is needed to efficiently use an L2. In the following,

---

[24] It is important to note that explicit and implicit *knowledge* and *learning* are not directly related, as explicit learning, which entails awareness on the part of the learner, can lead to both implicit and explicit knowledge (Elgort 2011). Muñoz (2013) explains that declarative knowledge can be regarded as a synonym for explicit knowledge, which means that this type of knowledge can be articulated. In contrast, implicit knowledge is related to procedural knowledge and knowing how to do something without being able to explain it.

evidence for such claims from coverage research, which also provides information on the aims of L2 vocabulary learning, is summarized. In a next step, findings on vocabulary learning outcomes in empirical studies are presented to see in how far learners achieve these aims in relation to vocabulary size.

In the past, L2 language learning has traditionally oriented itself at native speakers' language competence, which means that goals for vocabulary learning were frequently set in accordance with L1 users' vocabulary sizes. This approach is, however, questionable because a native speaker orientation in English language learning and teaching has generally been called into question (Seidlhofer 2001, 2011; Widdowson 1994) and because estimates of adult native speakers' vocabulary sizes differ widely (Nation 2006, see also Brysbaert et al. 2016; Goulden, Nation & Read 1990).[25] A more feasible way to establish goals for L2 vocabulary learning is found in research on lexical coverage, which attempts to estimate how many words L2 users of English need to know to adequately comprehend spoken or written English texts. However, in order to determine the vocabulary size needed for adequate comprehension, one first needs to establish what level of comprehension is 'adequate'. In an early study on reading comprehension Laufer (1989: 321) suggested that 95% lexical coverage is needed for "reasonable reading comprehension". However, a study by Hu and Nation (2000) showed "around 98% of coverage may be needed for most learners to gain adequate comprehension". A newer study by Laufer and Ravenhorst-Kalovski (2010) supports these previous findings because their results indicate that when adequate reading comprehension is conceptualized as a completely independent activity, a lexical coverage of 98% is needed; but if some guidance is provided a coverage level of 95% is appropriate. In contrast, another recent study by Schmitt, Jiang and Grabe (2011) found a linear relationship between lexical coverage and reading comprehension with no evidence for a threshold effect. However, based on the previous studies, 95% and 98% levels of coverage have become established as conventional learning goals in vocabulary research on reading and were also transferred to research on listening (e.g. Nation 2013). Van Zeeland and Schmitt (2013b) empirically investigated whether such a transfer of lexical coverage figures from reading to listening is justified and found the demands of listening to be somewhat lower. The authors thus suggest that "95 per cent may be the best lexical coverage target for L2 listening comprehension of informal narratives" (Van Zeeland & Schmitt 2013b: 474).

Based on the established levels of 95% and 98% lexical coverage needed for adequate comprehension, researchers have investigated various types of input (e.g. Nation 2006; Tegge 2017; Webb & Macalister 2013; Webb & Rodgers 2009a) to establish vocabulary size targets for L2 English learners. Studies on lexical coverage typically follow the same procedure: first, a

---

[25] More reliable studies in the past suggested receptive vocabulary sizes of about 17,000 words for adult L1 speakers (Goulden, Nation & Read 1990; Zechmeister et al. 1995). Recent research in New Zealand (Coxhead, Nation & Sim 2015) showed that the mean receptive vocabulary sizes of secondary school students ranged between 10,800 word families for 13-year-olds and 13,360 for 17-year-olds; however, some of the 17-year-olds had vocabulary sizes of up to 17,000 word families. Milton and Treffers-Daller (2013) found that the average vocabulary size of students at UK universities is about 10,000 word families and Brysbaert et al. (2016: 1) estimate that "an average 20-year-old native speaker of American English knows 42,000 lemmas and 4,200 non-transparent multiword expressions, derived from 11,100 word families".

corpus of input material is established, which is then compared to frequency lists created on the basis of large corpora, most frequently the BNC/COCA lists by Paul Nation (2006, 2012a), to see at which frequency level 95% and 98% lexical coverage are achieved in the corpus under analysis.[26] The results of lexical coverage research are thus presented in relation to frequency bands, which are based on the counting unit of word families (see section 3.1.1).

Table 3.2 summarizes the results of several studies and gives the vocabulary sizes needed for 95% and 98% coverage for a variety of written and spoken text types. As can be seen a minimum vocabulary size of 2,000 word families is needed for comprehension of graded readers, whereas unadapted written texts require a minimum knowledge of the 4,000 most frequent word families. Spoken texts have slightly lower lexical demands with knowledge of 3,000 to 4,000 word families necessary to reach 95% for a variety of input. A first analysis of multimodal input in digital games by Rodgers, Heidt and Wood (2019) indicates that the lexical demands of games are in between written and spoken texts.

Furthermore, the lexical demands listed in Table 3.2 are not uniform across text types; for instance, Webb and Rodgers (2009a) found that smaller vocabularies were needed to reach adequate comprehension levels in horror movies, dramas and crime films than in war films and animated movies. Dang and Webb (2014) investigated the lexical profile of academic spoken English and also found differences according to genre with the subcorpora of social sciences and arts and humanities showing lower lexical demands than the subcorpora of physical sciences and life and medical sciences. These rather different examples suggest that there can be considerable variation in terms of lexical demands within text types.

What becomes clear from the research findings presented in Table 3.2 is that a large amount of vocabulary knowledge is needed for the comprehension of written and spoken English input. In this respect, criticisms have been voiced that the use of word families as a counting unit in this research area may lead to underestimating the size of the learning task, for instance, Schmitt et al. (2017) propose approximate replications of four influential studies and make a case for redoing much of the existing research using lemmas as counting units. Until such research becomes available, care needs to be taken by both researchers and practitioners not to misinterpret or underestimate the currently available figures because, clearly, learners should aim to develop large L2 vocabularies for successful language use.

---

[26] In addition, almost all coverage figures given in such studies also include proper nouns, which are assumed to be known due to their low learning burden (Hirsh & Nation 1992). Many researchers also include marginal words, i.e. exclamations, interjections, hesitation makers and similar discourse markers, in their counts because they are common in spoken English and thus can be assumed to be known (Nation 2006). For detailed information on what is included in the coverage figures given in the different studies see Table 3.2.

Text type	Example analysed	95% coverage	98% coverage	Study
Reading				
Graded readers	*The Picture of Dorian Gray*	2000	3000	Nation (2006)
Graded readers	33 graded readers from Oxford Bookworm series	2000 + PN, MW	3000 + PN, MW	Webb & Macalister (2013)
Short novels	*Alice in Wonderland, The Haunting, The Pearl*	-	5000	Hirsh & Nation (1992)
Newspaper articles	Section A LOB Corpus	4000 + PN	8000 + PN	Nation (2006)
Novels	*Lady Chatterley's Lover*	4000 + PN	9000 + PN	Nation (2006)
Children's literature	*School journal*	5000 + PN, MW	10000 + PN, MW	Webb & Macalister (2013)
Prose texts for teenagers	Press/Fiction section of Wellington Written Corpus	5000 + PN, MW	10000 + PN, MW	Webb & Macalister (2013)
Listening				
Spoken English	Wellington Spoken Corpus	3000 + PN	7000 + PN	Nation (2006)
Pop songs	Wellington Corpus of Popular Songs	3000 + PN, MW	6000 + PN, MW	Tegge (2017)
Movies	Corpus of 318 movies	3000 + PN, MW	6000 + PN, MW	Webb & Rodgers (2009a)
Children's movies	*Shrek*	4000 + PN	6000 + PN	Nation (2006)
TV programmes	Corpus of 88 TV programmes	3000 + PN, MW	7000 + PN, MW	Webb & Rodgers (2009b)
Academic spoken English	BASE corpus	4000 + PN, MW	8000 + PN, MW	Dang & Webb (2014)
TED talks	Corpus of 60 TED Talks	4000 + PN	9000 + PN	Coxhead & Walls (2012)
TED talks	Corpus of 400 TED Talks	4000 + PN, MW	8000 + PN, MW	Nurmukhamedov (2017)
Multimodal input				
Digital games		5000 + PN, MW	9000 + PN, MW	Rodgers, Heidt & Wood (2019)

Table 3.2: English vocabulary sizes (given in word families) needed for 95% or 98% coverage level in various text types. (PN = proper nouns, MW = marginal words).

Data from empirical studies can shed light on the question whether learners of English actually achieve the learning aims established in coverage research. A starting point is provided by Laufer (2000: 48), who offers an overview of vocabulary sizes identified in empirical studies in relation to hours of instruction. Table 3.3 below reproduces Laufer's findings: as can be seen, vocabulary size tends to increase with hours of instruction, but the relationship is not linear. However, the data summarized by Laufer (2000) are relatively old by now and there is no information on how researchers arrived at these vocabulary sizes.

Size estimate[27]	Country	Hours of Instruction[28]	Reference
2,000 2,300	Japan (EFL University)	800-1200	Shillaw (1995) Barrow, Nakanishi & Ishino (1999)
4,000	China (English majors)	1800-2400	Laufer (2001)
1,220	Indonesia (EFL University)	900	Nurweni & Read (1999)
2,000	Oman (EFL University)	1350+	Horst, Cobb & Meara (1998)
3,500	Israel (High school graduates)	1500	Laufer (1998)
1,000	France (High school)	400	Arnaud, Béjoint & Thoiron (1985)
1,680	Greece (Age 15, High school)	660	Milton & Meara (1998)
1,200	Germany, (Age 15 High school)	400	Milton & Meara (1998)

Table 3.3: English vocabulary size of EFL learners: table adapted from Laufer (2000: 48).

Since a more recent systematic review of research findings on vocabulary size is not available, Table 3.4 summarizes data from recent studies with an emphasis on learners of English in Europe and more specifically in Austria. These data can be used to contextualize the results of the present study presented in Chapter 6. While data on hours of instruction are not always available, only studies that provide some indication of language level in relation to the CEFR (Common European Framework of Reference, Council of Europe 2001), the length of instruction in years or the hours of instruction were included. In addition, Table 3.4 includes information on the measurement technique used because clearly results concerning vocabulary size are not only related to actual learning outcomes, but also depend on the method of measurement employed in the studies. In general, the studies summarized in Table 3.4 employed three test formats: matching tasks, multiple choice items or a Yes/No format (see section 3.2). Results are reported either as an estimate of total vocabulary size in the case of multiple choice and Yes/No tests, or as mastery in relation to the frequency bands tested in the case of the Vocabulary Levels Test.[29]

---

[27] Unfortunately, units of counting are not reported for the estimates.

[28] The data on hours of instruction were largely obtained by Laufer through personal communication with colleagues from the respective countries (Laufer 2000: 47)

[29] Different mastery levels were set on the VLT by different researchers: Platzer (2006) and Tschirner (2004) set mastery at 80% (= 25/30), whereas Henriksen (2008), Stæhr (2008, 2009) and Szudarksi and Carter (2016) set mastery at 85% (= 26/30) in line with Schmitt, Schmitt and Clapham (2001). For the VocabLab test used by Peters (2018) a mastery level of 85% (26/30) was assumed in line with the discussion in the article. Further information on the tests mentioned can be found in section 3.2.3.

Size estimate	Participants	Country	Level	Measurement	Reference
2K level mastered by 10% (*M* = 9.28/30)	274 6th grade learners in primary school	Spain	6 YoI[30] (= 629 hours)	VLT	Augustín Llach & Terrazas Gallego (2009)
2K level mastered by 17% (*M* = 33.79/120 overall)	29 7th grade learners in a comprehensive school	Denmark	3 YoI	VLT	Henriksen (2008)
2K level mastered (*M* = 88.09/120 overall)	47 learners in 4th year of secondary school	Flanders, Belgium	3 YoI	VocabLab test	Peters (2018)
2K level mastered by 14%, 3K level by 9%	88 9th grade learners at the end of lower secondary	Denmark	7 YoI (< 570 hours)	VLT	Stæhr (2008)
2K level not mastered (*M* = 18/30)	41 learners in last grade of secondary	Poland	6 YoI	VLT	Szudarski & Carter (2016)
6604	93 8th grade learners at the end of lower secondary	**Austria**	A2+, 8 YoI (680 hours)[31]	VST	Zichtl (2017)
2K level mastered by 62%, 3K or 5K by 21% (*M* = 71.86/120 overall)	29 10th grade learners in upper secondary	Denmark	6 YoI	VLT	Henriksen (2008)
3041	83 learners in 10th grade of vocational middle schools	**Austria**	A2+, 9 YoI (≈ 800 hours)	V_YesNo	Hahn (2017)
7690	95 10th grade learners in upper secondary	**Austria**	B1, 10 YoI (≈ 920 hours)	VST	Zichtl (2017)
8850	87 12th grade learners at the end of upper secondary	**Austria**	B2, 12 YoI (≈ 1150 hours)	VST	Zichtl (2017)
2K level mastered by 81.9%, 3K level by 27.3%, AWL mastered by 44.3%	271 first-year students at university	**Austria**	B2, > 12 YoI	VLT	Platzer (2006)
2K level mastered by 94%, 3K level by 59%, 5K level by 30%, AWL mastered by 76%	142 first-year students of English at university	Germany	8 YoI (≈ 1280 hours)	VLT	Tschirner (2004)
2K and 3K level mastered (*M* = 99.63/120 overall)	32 first-year university students	Flanders, Belgium	6 YoI	VocabLab test	Peters (2018)
3K level mastered by 78%, 5K level by 49%	115 first-year university students	Denmark	B2-C1, 8 YoI	VLT	Stæhr (2009)
2K level mastered by 100%, 3K by 72%, 5K by 34%, 10K by 17% (*M* = 94.79/120 overall)	first-year university students	Denmark	9 YoI (min.)	VLT	Henriksen (2008)
5922	121 learners in last year of upper secondary/at university	Russia	B1+/B2	VST	Elgort (2013)
8855	40 university students with L1 Dutch in BA or MA degree	Belgium	11.43 YoI (mean)	VST	Elgort (2018)

Table 3.4: Findings of receptive English vocabulary size in studies on European EFL learners

---

[30] YoI is an abbreviation for Years of instruction and 'hours' refers to classroom contact hours, which may however not last a full 60 minutes; for instance, school lessons in Austria have a duration of 50 minutes, whereas in Germany they are only 45 minutes long.

[31] Zichtl (2017) does not report hours of instruction, the figures given were calculated on the basis of the Austrian national curriculum for academic secondary schools (see also section 4.2.1).

As can be seen from Table 3.4, different studies have produced very different results. While it is plausible that high frequency vocabulary in the form of the 2,000 most common word families (Nation 2013) is unlikely to be known after only three years of study (Henriksen 2008), it is rather discouraging that it is not known after six or seven years of instruction in some other studies (Augustín Llach & Terrazas Gallego 2009; Stæhr 2009; Szudarski & Carter 2016).[32]

In comparison, the average result of 6,000 word families for Austrian learners after 8 years of instruction by Zichtl (2017) is astonishing and likely to be inflated. Zichtl (2017) administered the Vocabulary Size Test (VST, Nation & Beglar 2007), which uses a multiple choice format and appears to contain a high number of cognates for L1 speakers of German.[33] Hahn's (2017) average finding of approximately 3,000 lemmas for vocational school students after 9 years of instruction thus seems more plausible. In general, there appear to be marked differences in relation to measurement: estimates based on the VST tend to be higher than those based on the Vocabulary Levels Test (VLT, Nation 1990; Schmitt, Schmitt & Clapham 2001), which emphasizes the importance of taking the test method into account when making cross-study comparisons.

Overall, it appears that the majority of European learners have mastered the 2,000 most frequent words by the end of lower secondary school and some reach the 3,000 word level at some point during upper secondary school (Elgort 2013; Henriksen 2008; Hahn 2017; Zichtl 2017). By the time learners have sat their A-levels and arrive at university, most of them know the 3,000 most common English words (Henriksen 2008; Peters 2018; Stæhr 2009; Tschirner 2004) and have substantial knowledge of academic vocabulary (Platzer 2006; Tschirner 2004). However, results for university students also differ widely between less than 3,000 (Platzer 2006) to more than 8,000 (Elgort 2018) after six to eleven years of instruction.[34]

Some of the studies in Table 3.4 refer to CEFR levels, since many European countries specify curricular learning goals in relation to this highly influential EU document. Although certainly not without critique (e.g. Hulstijn 2007), the CEFR is well-known in Europe among language teachers and learners as well as the general public. Hence, some attempts have been made to relate vocabulary knowledge to CEFR levels (Capel 2010, 2012; Meara & Milton 2003). Using data collected with the help of X-Lex from Greek and Hungarian learners, Milton & Alexiou (2009) found that learners need to acquire between 500 and 1,250 words to reach the next CEFR level with higher gains necessary for progression through the lower levels. However, their results are questionable because X-Lex only tests the 1K to 5K level; hence, the presence of a ceiling effect is very likely.

---

[32] In contrast, the learners in Peters (2018) had already mastered the 2,000 word frequency level after only three years of formal instruction, but this is likely due to the effect of extramural English to some extent: Flemish learners only start studying English in secondary school as their second foreign language after French, but they are exposed to English from an early age (De Wilde & Eyckmans 2017).

[33] I tried out the VST during piloting testing for the present study and found inflated results in comparison to the V_YesNo test. A likely reason is the high number of cognates as identified by two Austrian teachers of English (for the influence of cognates see also sections 3.1.3 and 3.2.2).

[34] For studies on the rate of learning and the development of vocabulary size over time see, for instance, Dóczi and Kormos (2016), Ozturk (2015, 2016) or Webb and Chang (2012).

The research presented so far has been concerned with receptive knowledge of written word forms, which is by far the most common test construct (e.g. Daller, Milton & Treffers-Daller 2007; Milton, Alexiou & Mattheoudakis 2014, see also section 3.2.1). Some studies have attempted to measure knowledge of spoken word form in recent years and most of them found that learners know fewer phonological than orthographic word forms (Cheng & Matthews 2018; Milton & Hopkins 2006; Milton, Wade & Hopkins 2010). However, some of these results have been called into question by Van Zeeland (2013, as cited in Schmitt 2014: 926–927), who used a more in-depth test format to probe knowledge of spoken word forms and conducted one-on-one interviews with her participants from different L1 backgrounds. Her results do not show the same advantage for orthographic over phonological knowledge as found in previous studies.

Another area that is not as well researched as receptive (orthographic) vocabulary size is productive vocabulary size, which is in part due to the fact that many productive test formats do not produce size estimates (see section 3.2.3). The most well-known productive size test is the Productive Vocabulary Levels Test (PVLT, Laufer & Nation 1999), which is the counterpart of the receptive levels test.[35] Since this study uses Lex30, a test that does not produce a concrete size estimate (see section 5.3.3.3) a detailed review is beyond the scope of this chapter. Table 3.5 limits itself to presenting estimates from a number of influential studies which have used the PVLT, although some of the data is relatively old by now.

In comparison to the receptive vocabulary size results as measured by the VLT and reported in Table 3.4, Table 3.5 shows notably lower figures for productive vocabulary size. Neither the university students in Waring's (1997) study, nor the learners at the end of high school or beginning of university in Laufer and Paribakht's (1998) study had mastered the high frequency vocabulary consisting of the 2,000 most frequent English word families. Laufer and Nation's (1999) more advanced participants had reached the 2K level, but unfortunately they provide little information about them. Finally, in Tschirner's (2004) study only little more than half of the university students had a productive vocabulary size larger than the 2,000 most frequent words.

In sum, this review of research on L2 English learners' achievement in relation to vocabulary acquisition again highlights that learners' receptive vocabulary is larger than their productive vocabulary. In addition, some studies suggest that knowledge of written word forms tends to be greater than knowledge of spoken word forms, but further research is needed in this respect. Research on vocabulary learning outcomes in Austria and Europe provides some context to the results of the present study; summarizing very broadly, European learners of English generally know the 2,000 most frequent words by the end of lower secondary school and the 3,000 most frequent words at university level in addition to academic vocabulary.

---

[35] It has to be pointed out that the cued gap-fill format of the PVLT has been criticized for not measuring truly productive knowledge, some scholars such as Read (2000: 126), even consider the PVLT an alternative form of receptive assessment rather than a productive measure.

Size estimate	Participants	Country	Level	Measurement	Reference
Have not mastered[36] 2K level (*M* = 11.8/18)	24 10th grade learners in high school	EFL environment	5 YoI	PVLT	Laufer & Nation (1999)
Have almost mastered 2K level (*M* = 15.0/18)	23 11th grade learners in high school	EFL environment	6 YoI	PVLT	Laufer & Nation (1999)
2K level mastered (*M* = 16.2/18), 3K level not mastered (*M* = 10.8/18),	18 12th grade learners in high school	EFL environment	7 YoI	PVLT	Laufer & Nation (1999)
2K level mastered (*M* = 17.0/18), 3K level not mastered (*M* = 14.9/18),	14 first-year students of English at university	EFL environment	-	PVLT	Laufer & Nation (1999)
Have not mastered 2K level (*M* = 14.9/18)	52 10th and 11th grade learners in high school and 27 first-year students of English at university	Israel	6-8 YoI	PVLT	Laufer & Paribakht (1998)
Have not mastered 2K level (*M* = 8.2/18)	76 first-year or second-year students of English at university	Japan	-	PVLT	Waring (1997)
2K level mastered by 53%, 3K level mastered by 8%	142 first-year students of English at university	Germany	8 YoI (≈ 1280 hours)	VLT	Tschirner (2004)

Table 3.5: Estimates of productive vocabulary size

3.1.3 Factors affecting vocabulary learning

Words differ in terms of their learning burden (Nation 1990, 2013), both in and of themselves and in relation to different learners. In addition, (incidental) vocabulary learning from input is also affected by characteristics of the context in which a new lexical item is encountered and by the level of processing that occurs during such encounters. Hence, this section will briefly discuss how lexical factors, input factors, learner factors and processing factors, summarized in Table 3.6, influence L2 (English) vocabulary learning and retention.

In terms of formal properties words are easier to learn if they are short, morphologically transparent and have a distinctive word form that is not easily confused with similar forms. In addition, pronounceability, phonotactic and orthographic regularity, familiarity, and regularity of phoneme to grapheme mappings also facilitate acquisition (Ellis & Beaton 1993; Ellis 1999; Laufer 1997b). With regard to grammatical features, reviews agree that nouns are easiest to learn, and adverbs are usually regarded as the most difficult word class, while the evidence for

---

[36] The level of mastery was set at 85% by Laufer and Nation (1999) and assumed to be the same for Laufer and Paribakht (1998), whereas Tschirner (2004) used a level of mastery of 80%.

verbs and adjectives, which lie somewhere in between, is inconclusive (Ellis & Beaton 1993; Laufer 1997b). Semantic factors also play a role: new L2 words are easier to acquire if the semantic distinction between concepts is drawn similarly in the L1 and the L2 (Ellis & Beaton 1993). In addition, form-meaning relations can facilitate learning in the case of onomatopoeia, but more often than not they emerge as complicating factors when multiple meanings are mapped onto one word form, as is the case with polysemous or homonymous words (Ellis 1999). Idiomaticity is another factor that can increase the learning burden, particularly in the case of multiword items, which often have idiomatic meanings that are not immediately transparent and cannot be inferred from the meanings of the constituent parts (Schmitt 2010: 53). In contrast, concreteness and imageability usually have a facilitative effect on learning (De Groot 2006; Elgort & Warren 2014; Ellis 1999; Ellis & Beaton 1993).

Lexical factors	Input factors	Learner factors	Processing factors
Word length	Frequency of occurrence	L1 background	Involvement load
Distinctiveness of word form	Salience	Existing L2 knowledge	Engagement
Pronounceability	Availability of contextual cues	Background knowledge	
Phonotactic regularity	Input complexity	Working memory	
Orthographic script		Gender	
Orthographic regularity		Aptitude	
Phoneme-grapheme correspondence		Motivation	
Part of speech		Vocabulary learning strategies	
Semantic content (in relation to L1)			
Correlation between form and meaning (onomatopoeia)			
Polysemy / homonymy			
Idiomaticity			
Concreteness and imageability of concept			
Frequency			

Table 3.6: Factors affecting vocabulary learning and retention

Probably the most important lexical factor is frequency, i.e. a given word's overall frequency in a language, since it affects most or all aspects of lexical processing and acquisition (Schmitt 2010: 13). Lexical items, including multiword items, with a higher frequency are learned earlier (see section 3.1.2) and they are also processed and remembered better (Ellis 2002; Schmitt 2010). Despite the all-pervasive importance of frequency, its measurement remains difficult, particularly in relation to L2 learning. Commonly, frequency information is drawn from large corpora, which provide a good indication of frequency in the language overall. However, these corpora only provide a reasonable approximation of L1 input – frequency in an L2 learner's

input, for instance through classroom discourse, may be very different. Thus, caution is needed when applying frequency information to L2 contexts or individual learners because "no corpus can replicate the exposure any individual person has, especially L2 learners" (González-Fernández & Schmitt 2015: 95).

Concerning input factors, frequency again plays a decisive role, although here it refers to frequency of occurrence in a given text, that is how often a word form is repeated (Ellis 1999). As Webb (2014: 1) states in a review on the role of repetition in incidental vocabulary learning, "[g]enerally, the more frequently words are encountered, the more likely they are to be learned", although a number of other lexical factors or input factors may influence the number of encounters needed to learn a novel lexical item. Findings from reading research indicate "that 8–10 exposures is the point where incidental learning begins to reach a critical mass and learning accelerates" (Pellicer-Sánchez & Schmitt 2010: 44, see also Malone 2018; Webb 2007), but results vary depending on the word knowledge aspects measured. Studies on incidental learning from auditive and audiovisual input (Peters, Heynen & Puimège 2016; Peters & Webb 2018; Van Zeeland & Schmitt 2013a; Vidal 2011) also show an effect for frequency of occurrence, but it is generally weaker than in studies on reading and more encounters are needed to develop different word knowledge aspects.

Further variables which affect vocabulary learning from input are a lexical item's salience, the availability of contextual cues, and the overall complexity of the input. Salience, i.e. the noticeability of a word in input (Ellis 1999), can be the effect of intrinsic lexical properties or it can be achieved, for instance, through stress in pronunciation, topicalization through syntactic structures, or by being essential to understanding the content. Similar concepts used in empirical studies are keyness (Elgort & Warren 2014) or word relevance (Peters & Webb 2018), but to date research has paid relatively little attention to the relevance of new words in input (see also Peters & Webb 2018). First results show a positive effect of salience on measures of explicit and implicit lexical knowledge in a study on incidental vocabulary acquisition from reading (Elgort and Warren 2014), whereas a study on vocabulary learning from watching television found no statistically significant effect (Peters & Webb 2018).

Contextual cues are a valuable resource for comprehension and are essential for lexical inferencing, which often is a first step in vocabulary learning from reading or listening (Laufer 1997a; Qian 2005). Learners' use of contextual cues has been studied extensively in research on reading (Bengeleil & Paribakht 2004; Paribakht 2005; Qian 2005; Webb 2008; Wesche & Paribakht 2010) and more recently also for listening (Van Zeeland 2014, see also sections 3.3.2.1 and 3.3.2.2). Results shows that learners generally pay more attention to local clues that occur in the same sentence as the unknown lexical item, or in the sentences immediately before and after, than to global clues that appear in the wider context of a given text. In addition, lexical inferences are generally more accurate when based on local clues in both reading (Qian 2005) and listening (Van Zeeland 2014).

Lastly, input complexity, that is the overall difficulty of a text in which new lexical items are encountered, evidently has an impact on vocabulary learning, for instance by affecting the comprehension of contextual cues available in the text. Ellis (1999) relates input complexity to lexical density defined as the ratio of content words to total word count and the density of unknown words. In the latter sense, input complexity is clearly dependent on learner's existing vocabulary knowledge and thus related to issues of lexical coverage. As discussed in section 3.1.2, research has shown that learners need high levels of vocabulary knowledge to achieve adequate comprehension of different types of texts (Nation 2006) and to allow the inference of unknown lexical items (Wesche & Paribakht 2010).

Learner factors also play a decisive role in the acquisition of new L2 vocabulary with the learners' L1 background being the most obvious influencing factor. As mentioned above, formal similarities or dissimilarities between L1 and L2 can impact vocabulary learning, particularly since psycholinguistic evidence shows that "the L1 is active during L2 lexical processing in both beginning and more advanced learners" (Schmitt 2010: 25–26). Another way in which the L1 background influences L2 vocabulary knowledge is the presence of cognates, which Elgort (2013: 255) defines as "words that are orthographically, phonologically and semantically similar across the two languages".[37] Although cognates may undergo some phonological and/or orthographic adaptation, they are typically recognizable for learners and have a considerably lower learning burden. Several studies including English and a variety of other languages (Elgort 2013; Laufer & McLean 2016; Laufer & Levitzky-Aviad 2018; Petrescu, Helms-Park & Dronjic 2017) show the presence of a cognate facilitation effect, which affects the scores of vocabulary size tests such as the VLT and the VST (see section 3.2.2). Due to the special status of English as the currently dominating lingua franca, many languages include English loanwords, which further increases the impact of cognates on the vocabulary knowledge of L2 learners of English. In addition to English loanwords in the L1, knowledge of further languages may also have a beneficial effect on vocabulary knowledge, as it increases the possibilities for crosslinguistic transfer.

A second influencing factor is learners' overall L2 proficiency and existing vocabulary knowledge (Ellis 1999). This aspect relates to input complexity and lexical coverage (see section 3.1.2): if too much vocabulary is unknown in a given text, the acquisition of new lexical items becomes very difficult or virtually impossible. There is ample evidence in research that vocabulary size has a positive effect on lexical inferencing from reading (Albrechtsen, Haastrup & Henriksen 2008, see also Wesche & Paribakht 2010 for a review) and listening (Van Zeeland 2014). Prior vocabulary also affects the retention rates of new lexical items acquired through reading (Pulido 2003, 2007; Webb & Chang 2015). Hence, (incidental) vocabulary learning from

---

[37] Cognates can be defined in different ways, but in vocabulary studies learners' ability to recognize a word as a cognate is more important than etymological criteria (Laufer & McLean 2016). Thus, Elgort (2013: 255) states that "[u]nlike the more restrictive understanding of *cognates* in linguistics as words of common etymological origin, in this study [and in other vocabulary studies] loan words or lexical borrowings [...] are classified as cognates, alongside words that have a common ancestor, as long as they match the criterion of having similar form and meaning across the L1 and L2".

input seems to conform to the Matthew principle because the greater a learner's vocabulary knowledge, and especially the larger their vocabulary size, the easier it becomes to acquire new words. A further aspect that can affect L2 vocabulary acquisition from reading or listening is learners' background or world knowledge. Several studies (see Van Zeeland 2014: 1008 for a review) have shown that background knowledge is used in lexical inferencing processes and that topic familiarity has positive effects on lexical inferencing success and the acquisition of word knowledge in both reading (Pulido 2003, 2007) and listening (Van Zeeland 2014).

Learners' attentional and cognitive resources also contribute to successful vocabulary acquisition (Elgort & Nation 2010). As Pavičić Takač (2008: 10) puts it "[t]he role of memory is crucial in any kind of learning and vocabulary learning is no exception". Research has established a link between phonological short-term memory and the learning of new words for both L1 and L2 vocabulary acquisition and has found that working memory capacity affects the noticing of new lexical items as well as their retention (Dóczi & Kormos 2016). While there have been studies on the role of working memory in intentional word learning, Malone (2018) is the first study to explore the role of working memory in incidental vocabulary learning from reading.

With the exception of vocabulary learning strategies (VLS), other learner-related factors, or individual differences (Dörnyei 2005; Dörnyei & Skehan 2003), such as gender, age, aptitude or motivation have not been studied extensively in relation to L2 vocabulary acquisition.[38] Vocabulary learning strategies have been researched since the late 1980s (Gu 2013) and in the 1990s VLS taxonomies were constructed in line with tendencies in the broader field of strategy research. The two most widely known examples are by Gu and Johnson (1996) and Schmitt (1997), who differentiates between discovery and consolidation strategies. Schmitt (1997) found that in a sample of 600 Japanese EFL learners the most frequently reported discovery strategies were using bilingual dictionaries, guessing from textual context and asking classmates for help, while the most frequent consolidation strategies were verbal repetition, written repetition, studying the spelling and sound of a word, saying new words out loud, taking notes in class and using word lists. Learner strategies in general and VLS in particular were widely researched in the 1990s (see the edited volume by Cohen & Macaro 2007 for a review), but most studies focused on instructed contexts. The lack of studies on independent settings has been criticized by Klapper (2008) because "a lot of vocabulary is, after all, learnt independently of the classroom" (Klapper 2008: 161). A recent study by Bytheway (2015), which qualitatively investigated learners' use of VLS while playing a massively multiplayer online role-playing game (MMORPG), seems to be the first study that actually examined the role of VLS in an extramural context.

---

[38] An exception is the recent edited volume by Jiménez Catalán (2010), which presents gender perspectives on second and foreign language vocabulary acquisition. There is also a suite of tests concerned with language learning aptitude by Meara (2005): the LLAMA tests are loosely based on the MLAT (Carroll & Sapon 1959) and have been the subject of recent validation studies (Granena 2013; Rogers et al. 2016; Rogers et al. 2017). In relation to vocabulary acquisition, the recently revised LLAMA B (Meara & Miralpeix 2017), which tests paired associate learning, is of particular interest, but at present it is not clear how exactly scores are related to vocabulary learning ability.

The lack of studies on independent or informal settings may also be a consequence of the general criticism of learner strategy research, which has been mounting in the last 15 years (Dörnyei & Skehan 2003, see also Rose et al. 2018). Dörnyei (2005) proposed using the notion of self-regulation instead of learner strategies. A new model of self-regulatory capacity, a concept taken from educational psychology, was first presented and applied to vocabulary learning in a seminal study by Tseng, Dörnyei and Schmitt (2006). This study also presents a new measure, the Self-Regulating Capacity in Vocabulary Learning (SRCvoc) scale, which was meant to replace previous VLS instruments as these were criticized for not being psychometrically valid. The SRCvoc operationalizes self-regulatory capacity as five facets: commitment control, relating to commitment to goals; metacognitive control, relating to concentration and monitoring; satiation control, relating to boredom or interest in a task; emotion control, relating to emotional management; and environmental control, relating to the management of environmental influences on the learning process (Tseng, Dörnyei & Schmitt 2006). Although initial validation evidence was good, the SRCvoc has since been used in validation studies in several contexts (Bilican & Yeşilbursa 2015; Doaee, Sarkeshikian & Tabatabaee 2017; Mizumoto & Takeuchi 2012; Sarkeshikian, Tabatabaee & Doaee 2018; Yeşilbursa & Bilican 2013) which provided mixed results and indicate a need to further investigate the scale.

While motivation and its impact on language learning have been extensively researched in the last decades (Al-Hoorie 2017; Dörnyei & Ryan 2015; Dörnyei & Ushioda 2009), it has not been specifically looked at in the area of L2 vocabulary learning (Laufer & Hulstijn 2001). Nevertheless, "it is logical to assume that motivation also facilitates vocabulary learning" (Tseng & Schmitt 2008) because of evidence from the wider field of SLA. Building on Tseng, Dörnyei and Schmitt (2006), Tseng and Schmitt (2008) is the first study to explicitly link vocabulary acquisition and motivation in a model of motivated vocabulary learning.[39] Their hypothesized model consists of six parts, which were operationalized in a measurement instrument and administered to 250 participants from China and Taiwan. Analysis of the empirical data using structural equation modelling resulted in a cyclical model, Tseng and Schmitt (2008) interpret this finding as suggesting that vocabulary learning is a systematic, yet dynamic and cyclical process. Because vocabulary acquisition is an ongoing process (see section 3.1.2), it would be ideal if learners developed into autonomous, self-motivated experts who ideally display great capacities for emotional management, self-regulatory capacity, strategy use, critical awareness of learning and understanding of vocabulary development; however, as Schmitt and Tseng (2008) acknowledge such expertise is far from easily achieved. Still, their study provided an important starting point for further research on motivation as an influencing factor on vocabulary learning as indicated by three more recent studies (Tanaka 2017; Zhang et al. 2017; Zheng 2012).

---

[39] A detailed explanation is not possible here for reasons of scope, see Schmitt (2010: 94–97) for a broad overview and Tseng & Schmitt (2008) for detailed information.

Finally, vocabulary learning and retention are also affected by the level of processing a lexical item receives upon noticing it. Since lexical processing is a topic of its own (see Barcroft 2015) it cannot be covered in detail and depth of processing is included as one influencing factor here. As Schmitt (2010: 26) puts it, "[i]t is a commonsense notion that the more a learner engages with a new word, the more likely he/she is to learn it". The first proposal which put this notion into a coherent theory was the *Depth of Processing Hypothesis* by Craik and Lockhart (1972). Drawing on their work as well as further developments in the wider field of SLA, Laufer and Hulstijn (2001) proposed the concept of task-induced involvement, which has become known as the *Involvement Load Hypothesis*. They conceptualize task-induced involvement as a motivational-cognitive construct consisting of three factors*:* need, a motivational component; search, a cognitive component; and evaluation, a second cognitive component.[40] Laufer and Hulstijn (2001: 15) further posit that all "[o]ther factors being equal, words which are processed with higher involvement load will be retained better than words which are processed with lower involvement load."

Empirical evidence (Eckerth & Tavakoli 2012; Hulstijn & Laufer 2001; Keating 2008; Kim 2008; Nassaji & Hu 2012; Tahmasbi & Farvardin 2017; Zou 2017) seems to generally support the concept of task-induced involvement as proposed by Laufer and Hulstijn (2001), although it has also been subject to criticism. For instance, Schmitt (2008) argues that the Involvement Load Hypothesis does not fully take the role of the learner into account, for example, in relation to self-regulatory and strategic behaviour, and that many more factors than those captured by Laufer and Hulstijn's proposal affect vocabulary learning. He thus proposes the term *engagement* as a cover term to capture all possible forms of involvement in vocabulary learning (see also Schmitt 2010: 26–28). Nation and Webb (2011) voice a similar criticism and suggest a new approach called *Technique Feature Analysis*, which they present in form of a checklist to be used for assessing the effectiveness of vocabulary learning tasks. It includes aspects in the five areas of motivation, noticing, retrieval, creative use and retention. Using several examples, Nation and Webb (2011) show that the assessment of tasks according to the Involvement Load Hypothesis and Technique Feature Analysis does not necessarily lead to the same results, but since both proposals can be evaluated empirically, they can be tested against each other in practice. To date, only two studies have compared the explanatory value of the Involvement Load Hypothesis to that of Technique Feature Analysis: Hu and Nassaji (2016) and Gohar, Rahmanian and Soleimani (2018) both compared performance on tasks rated differently on the two scales and found that overall Technique Feature Analysis accounted for more variance and had higher predictive power.

---

[40] *Need* as the motivational component refers to a need to fulfil a certain task based on inner drive or external requirements, but Laufer and Hulstijn (2001) assume that need is stronger when it is intrinsic to the learner. *Search* is a cognitive component relating to attempts "to find the meaning of an unknown L2 word or trying to find the L2 word form expressing a concept" (Laufer & Hulstijn 2001: 14), thus essentially finding the information that is needed, and *evaluation* refers to the cognitive process of assessing the adequacy of a lexical item in a given context.

It has become clear that vocabulary learning is a multifaceted process which is influenced by a multitude of factors. This overview has focused on vocabulary acquisition from input as this form of learning is most relevant for the empirical study presented in Chapter 5, but it is obvious that especially in instructed contexts other factors such as the intention to learn, an external requirement to memorize a lexical item such as a test, or manipulation of lexical items and their properties (Schmitt 2008), which were not discussed here, also play a role.

This section presented the foundations of vocabulary research and findings relating to vocabulary learning aims, processes and outcomes. We have seen that despite the importance of these constructs for empirical research there are no universally agreed upon conceptualizations of what counts as a 'word' and what it means to 'know a word', so that these issues should ideally be addressed in every empirical study. Concerning L2 vocabulary development, the incremental nature of the learning process and the role of frequency have been highlighted; in addition, we have seen that receptive knowledge generally precedes productive knowledge and that, like every other type of learning, vocabulary acquisition also involves attrition phenomena. In relation to the aims of vocabulary development, research on lexical coverage suggests that learners need a minimum of 2,000 word families to deal with graded readers and more than 4,000 word families to read authentic texts. For spoken input, a minimum of 3,000 word families appears to be necessary to understand pop songs, movies or TV programmes. In order to achieve these goals, both incidental and intentional learning are useful: whereas intentional learning leads to greater gains in the short term, it seems impossible that all lexical items needed for successful L2 comprehension and use can be acquired intentionally; clearly, incidental learning from input is needed as well. A review of recent findings concerning vocabulary size shows that vocabulary learning is a slow process: by the end of lower secondary school, most European learners have mastered the 2,000 most frequent word families as measured by the VLT and VST and reach the 3K level during upper secondary school. However, there are great differences in terms of vocabulary size depending on the learning context and the test used. In addition, a variety of lexical, input-related, learner-related and processing factors also influence learning processes and achievements in L2 vocabulary development.

3.2   Measuring vocabulary size

While this study is not centrally concerned with vocabulary measurement, vocabulary tests and related issues need to be discussed because methods of measurement have a large impact on empirical results, as became evident, for instance, in relation to vocabulary size in section 3.1.2. It is vital to take this seemingly obvious fact into consideration in SLA studies focusing on vocabulary because "[i]n spite of the very long history of interest in vocabulary measurement, there are remarkably few standard, well-researched vocabulary tests" (Laufer & Nation 2012: 165). Kremmel (2017) also critically comments on the fact that validation evidence is scarce even for the most widely used vocabulary tests, which leads to problems when it comes to the validity and comparability of empirical findings. This section briefly summarizes issues concerning

construct definitions and their operationalization (3.2.1) and the related aspects of validity, reliability and usefulness with a focus on vocabulary size (3.2.2), before providing an overview of the most commonly used size tests (3.2.3).

3.2.1 Defining and operationalizing the construct

Like all other language tests, vocabulary tests are assumed to measure a construct, that is an "underlying ability or mental attribute" (Read 2000: 8). In the case of vocabulary testing the process of construct definition relates to the conceptualizations of a word and word knowledge. As discussed in section 3.1.1, several units of counting and a variety of different word knowledge conceptualizations have been proposed and thus construct definitions of available vocabulary tests are also highly diverse.

Building on earlier work by Chapelle (1998) and Read (2000), Read and Chapelle (2001) distinguish between three approaches to construct definition in vocabulary testing: trait definitions, behaviourist definitions and interactionalist definitions. They criticize that

> there has been no real tradition in vocabulary testing of construct definition in any explicit form. Implicitly, this area has been dominated by trait definitions, operationalized in discrete, selective and context-independent tests of learners' knowledge of individual words presented in isolation (Read & Chapelle 2001: 9).

In contrast to behaviourist and interactionalist definitions, trait definitions regard L2 users' vocabulary knowledge as a personal trait that can be measured in isolation and "without reference to any particular context of use" (Read & Chapelle 2001: 8). Vocabulary size tests are typical examples of such a construct definition as they present target items in isolation and assume that a learner's test performance reflects their underlying vocabulary knowledge. Read and Chapelle's (2001) criticism is still valid because even newer tests, such as the VST or the updated VLT (see section 3.2.3), rely on trait definitions and conceptualize vocabulary as a construct that depends on individual characteristics and can be measured independent of context. While such tests are useful because many different learners can be compared, context still needs to be taken into consideration when interpreting their results.[41]

The use of trait definitions, however, does not indicate the type of word knowledge tested, which is closely linked to test purpose.[42] Vocabulary size tests typically measure knowledge of the form-meaning link (Schmitt 2014), but as mentioned in section 3.1.2 the form-meaning connection can be mastered at different levels. Drawing on Laufer and Goldstein's (2004) study, Schmitt (2010: 86) distinguishes between the two aspects of form and meaning and the two levels of mastery of recognition and recall, resulting in the four options summarized in Figure 3.1:

---

[41] See for instance the discussion of cognates in section 3.2.2 as one example of how context, more specifically test takers' L1, should be taken into account.

[42] Test purpose also includes aspects other than the type of word knowledge tested. Following Read and Chapelle (2001), it consists of three components: inferences to be drawn from the test performance, which includes decisions on the type of knowledge tested; uses of the test results, and intended impact.

	Degree of mastery	
	Recognition	*Recall*
Meaning	Meaning recognition (select definition, L1 translation, etc.)	Meaning recall (supply definition, L1 translation, etc.)
Form	Form recognition (select the L2 item)	Form recall (supply the L2 item)

Figure 3.1: Four constructs in relation to the form-meaning link based on Schmitt (2010: 86)

If receptive vocabulary size related to the activities of reading and listening is being targeted, then test designers will want to test meaning recall; however, if L2 users' vocabulary knowledge in relation to language production is of interest, form recall should be tested (see also Read 2000: 154–157). As Schmitt (2010: 87) points out, form and meaning recognition knowledge "probably only come into play in reference look-up situations in the real wold", but they are useful measurement constructs to tap into the initial stages of word learning.

Once the construct has been defined it needs to be operationalized, which involves decisions about the sampling of target items, item formats, scoring and the presentation of results. A useful starting point for a discussion of these issues in relation to existing vocabulary tests is Read's (2000) framework of vocabulary assessment, which identifies three dimensions. The first of Read's (2000: 9) dimensions is a continuum of discrete to embedded measures and thus refers to construct definition, the second relates to the selection of lexical target items and presents a continuum of selective to comprehensive measures, and the third is a continuum ranging from context-independent to context-dependent measures and concerns the test format. As discussed in the previous section, vocabulary size tests are discrete measures in terms of construct. Furthermore, size tests are commonly selective in relation to sampling, and they tend to be context-independent with regard to test format, as target words are usually presented in isolation or short non-defining contexts.

Several decisions need to be taken when operationalizing the construct of a vocabulary size test. First, lexical target items need to be selected, since evidently not all words of any given language can be tested. In order to establish the sample to be tested, first the overall target population of words needs to be defined. In the past, dictionaries were frequently used as a reference source for the overall population (see Schmitt 2010), more recently frequency lists based on large corpora such as the BNC or COCA often serve as the basis of vocabulary size tests (see also section 3.1.1). Once the target population has been decided on, it is common for the test designer to "select a sample of these items to fix on a test, and then assume that the percentage of items answered correctly on the test represents the percentage of items known in the total population" (Schmitt 2010: 193).[43] A key consideration in the sampling process is the unit of counting, which has a large impact on the interpretation of results (see section 3.1.1). In addition, the overall

---

[43] For further discussion of lexicographical problems in relation to sampling and the assumptions underlying size tests see Meara (1996).

population is commonly split into frequency bands, akin to sampling frames, before selecting the target items for the test.[44] It has been a long-standing convention in the field to use 1,000 word frequency bands for the sampling of target items and the establishment of test levels. This convention has, however, been recently called into question by Kremmel (2016), who argues that it would be more informative for diagnostic purposes to use smaller frequency bands of 500 words at the high frequency levels up to the 3K level, then use conventional 1,000 word bands for mid-frequency vocabulary and frequency bands larger than 1,000 for low-frequency items. In addition, Kremmel's (2016) study indicates that in terms of coverage it may not be worthwhile to test words beyond the 10,000 word frequency band. Kremmel's suggestions, however, have yet to be taken up, so that almost all currently available size tests use 1,000 word frequency bands.

Having decided on an overall target population and sampling frames, the actual target items need to be selected. This selection process is determined by the sampling rate, i.e. the number of items to be selected from each frequency band, which in turn interacts with the chosen test format. Clearly, test formats differ in terms of the difficulty of writing test items, the time and effort needed for completion by test takers, and the amount of time needed for scoring. For instance, a multiple choice test requires much more time than a Yes/No checklist test both in terms of writing and completion because writers need to come up with plausible distractors and test takers need to read and process much more information. Hence, decisions on the number of items to be tested need to be considered in relation to the test method required by the construct definition. In sum, "[t]he choice of format certainly depends on the kind of information and the degree of precision of knowledge a test developer or user is aiming for" (Kremmel 2017: 29). Typical test formats include checklist formats, gap-fill or cloze tests, matching formats, multiple choice items and L1/L2 translations or L2 definitions, each with their own advantages and disadvantages in relation to item writing, test administration, scoring procedures and the interpretation of results (for a recent critique of several widely used test format see Kremmel & Schmitt 2016).[45]

Building on the discussion in section 3.1.1, this section briefly summarized approaches to construct definition in vocabulary measurement and related criticism before giving a very general account of the process of operationalizing a test construct. Vocabulary size tests, which are of most interest in relation to the present study, are commonly characterized by trait definitions, which assume that vocabulary knowledge can be tested independently of context as it is seen as a characteristic of the test taker. Typically, the construct of vocabulary size tests is specified in relation to the form-meaning link; thus, Schmitt's (2010) distinction between four

---

[44] Frequency is currently the most common clustering factor used in item sampling, but as Kremmel (2017: 26) argues it is "not a sufficient predictor of knowledge" on its own. Experimental studies may also control for additional features such as part of speech or word length, but this is not usually the case with size tests.
[45] A detailed review of test formats used in vocabulary measurement is well beyond the scope of this thesis, but see section 3.2.3 for a short overview of well-known vocabulary size tests and Beglar and Nation (2014), Schmitt (2010) and Kremmel (2017) for further information and an overview of test formats used in existing vocabulary measures.

different levels of knowledge of the form-meaning link is useful to categorize size tests in terms of the type of word knowledge they measure (see section 3.2.3). In addition, the short overview in relation to the operationalization of test construct highlights the range of the decisions that need to be taken with regard to the selection of target items as well as test format. More detailed information on the two vocabulary measures used in the empirical study, including an in-depth discussion of their advantages and drawbacks, will be provided in Chapter 5.

3.2.2 Issues of validity, reliability and usefulness

As has been mentioned at the beginning of this section, a plethora of different vocabulary measures have been proposed, but little evidence is available to back up the claims made for most of these tests. This situation has been severely criticized by Kremmel (2017: 1–2):

> The field of vocabulary assessment seems notorious for a cottage-industry mindset, in which validation evidence is sparse for even the most prominent and most used vocabulary tests, and in which mere assumptions have become unquestioned traditionalized conventions and any "new" vocabulary test seems just another ostinato.

The lack of validation evidence has consequences for the use of tests both as research and pedagogical tools: for many tests it is not clear whether they actually measure what they purport to measure, which affects their practical usefulness in applied research and instructional settings. While new theories of validity and corresponding procedures, such as argument-based validation (Kane 1992, 2012) or the socio-cognitive approach (O'Sullivan & Weir 2011; Weir 2005), are widely discussed in the field of language testing, their impact on the field of vocabulary testing has been limited. Voss (2012) and Kremmel (2017) present two notable exceptions as both follow an argument-based validation process in relation to a productive test of academic collocational ability (Voss 2012) and a diagnostic and computer-adaptive profiler of vocabulary knowledge (Kremmel 2017).

In practice, vocabulary measures are often claimed to be valid instruments without much empirical evidence to support this claim. If evidence is available, it is frequently provided in the form of concurrent validity in relation to a second, 'established' vocabulary test, even though the constructs may not be congruent. However, like all other language tests, vocabulary tests are not valid instruments per se, since validity should not be regarded as an intrinsic property, but as an "argument concerning interpretation and use" (Brunfaut & Schmitt 2018: Slide 70) that needs to be established in relation to different contexts. Read and Chapelle (2001) argue that validation can only be carried out in relation to a test purpose (see also footnote 42 in section 3.2.1). A clear specification of test construct and test purpose is thus of central importance in order to collect validation evidence and counteract threats to validity.

One threat to validity that has received considerable attention in the field of vocabulary testing in recent years is the influence of learners' L1 background and the presence of cognates on test performance. As discussed in section 3.1.3, Elgort (2013), Laufer and McLean (2016), Petrescu, Helms-Park and Dronjic (2017), and Laufer and Levitzky-Aviad (2018) all found a significant facilitative effect of cognates on test performance, which raises the question whether cognates

should be included in vocabulary (size) tests. Laufer and McLean (2016) argue that from one perspective cognates or loanwords are a legitimate part of learners' vocabulary knowledge, but vocabulary size tests are taken by L2 users with many different L1 backgrounds, which leads to differing proportions of loanwords on the test for each test taker. As shown in their study, this situation may result in a "built-in bias that poses a threat to test validity" (Laufer & McLean 2016: 215) and thus size tests should ideally take L1 backgrounds into account, although this solution may be practically impossible. Elgort (2013) comes to a similar conclusion that the proportion of loanwords included in a vocabulary test should be similar to the overall proportion of loanwords in the target language and Petrescu, Helms-Park and Dronjic (2017) suggest that researchers may want to consider cognate and non-cognate items separately. However, Laufer and Levitzky-Aviad's (2018) findings indicate that random sampling of target words does not lead to the expected overestimation due to the cognate facilitation effect because test versions with a random and a representative number of loanwords produced very similar results. While this outcome is reassuring in relation to existing size tests, further research with participants from different L1 backgrounds is needed and the possibility of over- or underestimating test takers' vocabulary knowledge due to cognate effects needs to be kept in mind.

Two further issues that are part of the validation process but deserve to be mentioned in their own right are the reliability and usefulness of test scores. Similar to validity, reliability, i.e. the consistency of measurement (Bachman & Palmer 1996: 19), is not a characteristic of a test that remains constant once established, but a variable that should be investigated for each new context of use. Vocabulary researchers typically establish reliability of test scores through measures of internal consistency such as Cronbach's alpha, and few studies report evidence from more rigorous procedures such as test – re-test designs. More in-depth investigations of reliability, at least in relation to new instruments, thus remain a desideratum as well. In addition, some measures are limited in their practical usefulness.[46] Clearly, usefulness or utility needs to be considered in relation to test purpose, but at present relatively few vocabulary measures include comprehensive guidelines for the use and interpretation of test scores. Results of vocabulary size tests are typically presented as an estimated level or an estimate of total size for each test taker, but little guidance is given on what these scores mean in relation to actual language performance. This is particularly crucial from a pedagogical perspective, as language teachers cannot be expected to be experts of L2 vocabulary assessment. Moreover, there are tests whose scores can be used to compare different test takers in a given context, but little is known about what their scores mean in terms of participants' actual vocabulary knowledge.[47] Despite interesting methodological innovations, the usefulness of such vocabulary measures is necessarily limited. From the perspective of a vocabulary test user rather than designer usefulness in relation to one or more test purposes should thus be considered more centrally.

---

[46] Despite the similarity in name, I am not referring to the overall theoretical framework by Bachman and Palmer (1996) here, but a much more practical quality relating to what Read and Chapelle (2001: 16) call the utility of test scores.

[47] For want of a better alternative, one such test - Lex30 - was used in the present study and will therefore be discussed in more detail in Chapter 5.

As this short summary shows, measuring vocabulary knowledge in a valid, reliable and useful way is a difficult task. Concerns regarding methodological issues and the lack of validation evidence for vocabulary tests have been raised repeatedly, most recently by Kremmel (2017). Based on a much more detailed review he identifies six areas for improvement which relate to aspects of construct definition, operationalization and validation:

> In summary, it can thus be argued that currently existing vocabulary tests suffer from six major weaknesses: (1) focus on single words, (2) inappropriate sampling in terms of unit of counting, frequency bands and representativeness, (3) problematic or unprincipled selection of item formats, (4) favouring of written over spoken vocabulary knowledge, (5) focus on single dimensions of word knowledge, and (6) generally insufficient validity evidence (Kremmel 2017: 69–70).

Clearly, there is much room for improvement from a measurement perspective to which I might add a critique of usefulness from a user-oriented perspective. Currently, the field of vocabulary testing resembles a jungle that is hard to navigate for non-expert users of vocabulary tests, which is why the usefulness of tests should also be taken more centrally into account and more clearly communicated by test designers. At the same time, we have already accumulated a substantial body of knowledge about L2 vocabulary development with existing measures and it is to be hoped that the movement towards better test validation and more inclusive measures of vocabulary knowledge, including the newest suggestions concerning online measures of vocabulary processing and learning (Godfroid 2019), will help to advance lexical research as a whole.

3.2.3 Tests of vocabulary size

This section presents some of the most well-known vocabulary size tests as two such measures are employed in the current study. However, as Schmitt (2010: 216) points out, it is impossible to measure vocabulary size without assumptions about depth of knowledge:

> [A]ll size measures have a (sometimes implicit) criterion of minimum knowledge for a lexical item to be counted as 'known'. [...] Thus, it can be said that all size measures are also depth measures in the sense that some quality of knowledge, no matter how minimal, must be operationalized as the criterion of sufficient knowledge.

Most size tests measure knowledge of the form-meaning link, but as mentioned in section 3.2.1, the form-meaning link can be conceptualized at four different levels with regard to the aspects of form and meaning and the levels of mastery of recognition and recall. These also relate to depth as argued by Schmitt (2010) in the citation above (see also Schmitt 2014). In addition, size tests can also be classified according to written or spoken test modality. Table 3.7 includes all vocabulary size tests mentioned in this thesis organized according to the latter two criteria. In the following, these tests will briefly be discussed in turn. In addition to the tests mentioned in Table 3.7, empirical studies have also used translation formats (e.g. Stubbe 2013), both at the level of L2 to L1 recognition and L1 to L2 recall, and interview procedures (e.g. Schmitt, Schmitt & Clapham 2001; Van Zeeland 2013, as cited in Schmitt 2014: 926–927) to measure vocabulary size.

Test	Abbrevi-ation	Construct	Format	Results	Reference
Tests of written receptive vocabulary size					
LexTALE	-	written form recognition (lexical decision task)	checklist of 60 items including 20 pseudowords	raw score/ percentage score	Lemhöfer & Broersma (2012)
VocabLab test	-	written meaning recognition	five-option multiple choice format (incl. I don't know option) with 30 items per level	mastery of 2K, 3K, 4K, 5K level (lemmas)	Peters, Velghe & Van Rompaey (2015)
Vocabulary Levels Test	VLT	written form recognition	matching format with 3 target words and definitions and 3 distractors per item cluster, 30 items in 10 clusters per level	mastery of 2K, 3K, 5K, 10K level and academic vocabulary (word families)	Nation (1983, 1990); Schmitt, Schmitt & Clapham (2001)
Updated Vocabulary Levels Test	New VLT	written form recognition	matching format with 3 target words, 3 definitions and 3 distractors per item cluster, 30 items in 10 clusters per level	mastery of 1K, 2K, 3K, 4K, 5K level (word families)	Webb, Sasao & Ballance (2017)
Vocabulary Size Test	VST	written meaning recognition	four-option multiple choice format	estimate of 14,000 or 20,000 most frequent word families	Nation & Beglar (2007), Nation (2012b)
Yes/No tests, e.g. V_YesNo	-	written meaning recall (without evidence)/form recognition	checklist including pseudowords (V_YesNo: 100 target and 100 pseudowords)	estimate out of 10,000 most frequent lemmas	Meara (2015a), see also Meara (1992, 2010)
X-Lex	-	written meaning recall (without evidence)/form recognition	checklist with 120 items including 20 pseudowords	estimate of 5,000 most frequent lemmas	Meara & Milton (2003)
Tests of aural receptive vocabulary size					
Aural Lex	A-Lex	aural meaning recall (without evidence)/form recognition	checklist with recorded prompts and 120 items including 20 pseudowords	estimate of 5,000 most frequent lemmas	Milton & Hopkins (2005)
Listening vocabulary levels test	LVLT	aural meaning recognition	four-option multiple choice format with recorded promts and 24 items per level (30 for AWL)	mastery of 1K, 2K, 3K, 4K, 5K level and AWL (word families)	McLean, Kramer & Beglar (2015)
Peabody Picture Vocabulary Test (4th edition)	PPVT-IV	aural meaning recognition	aural prompts and picture cards, 12 sets of 12 items	raw score of max. 204 test items	Dunn & Dunn (2007)
Picture Vocabulary Size Test	PVST	aural meaning recognition	four-option picture multiple choice format with optional "I don't know" option and aural and written prompts, 96 items	estimate of 6,000 most frequent word families	Anthony & Nation (2017)

Test	Abbrevi-ation	Construct	Format	Results	Reference
Tests of written productive vocabulary size					
Lex30	-	written form production/recall	word association task with 30 cue words	raw score of max. 120 points	Meara & Fitzpatrick (2000)
Productive Vocabulary Levels Test	PVLT	written cued form recall	cued gap-fill task with 90 items across 5 levels	mastery of 2K, 3K, 5K, 10K level and academic vocabulary (word families)	Laufer & Nation (1999)
Tests of written vocabulary size and strength					
Computer Adaptive Test of Size and Strength	CATSS	written meaning recognition/recall and form recognition/recall	Computer adaptive test format across four test modalities with 150 target words (600 items in total)	mastery of 2K, 3K, 5K, 10K level and academic vocabulary (word families) in relation to four modalities	Laufer & Goldstein (2004)

Table 3.7: An overview of prominent tests of vocabulary size

Beginning with the simplest test format for receptive knowledge of written form, checklist or Yes/No tests, such as LexTALE, V_YesNo or X-Lex, present test takers with a list of target words for which they are asked to tick 'yes' if they know a word according to the test instructions or 'no' if they do not. To prevent test takers from guessing, the checklist also includes pseudowords and the score is corrected if these are selected as known. This simple format of Yes/No tests has several advantages: the test is easy to take, and little time is needed for administration, which also allows a large sample of words to be tested (Meara 1996; Meara & Miralpeix 2017; Schmitt 2010). One major drawback is, however, the inconclusive evidence for the best scoring method: several formulae for correcting the number of correct responses by the number of pseudowords ticked have been proposed over the years and there is no consensus on the best option (see section 5.3.5.3). Many Yes/No tests have been produced by Paul Meara and colleagues (e.g. Meara & Jones 1988; Meara 1992, 2015a); these tests measure meaning recall knowledge according to the instructions because participants are supposed to choose 'yes' if they know the meaning of a word, although they do not have to provide concrete evidence for their knowledge. In contrast, the Yes/No test designed by Lemhöfer and Broersma (2012), LexTALE, asks test takers to indicate whether word forms actually exist in English, and thus most closely resembles form recognition, although the form-meaning link is not explicitly tapped into.

One of the best-known sizes tests, the Vocabulary Levels Test (VLT, Nation 1983) is often regarded as the closest to an "industry standard" (Meara & Miralpeix 2017: 114). It uses a matching format that appears relatively complicated at first glance: test takers are presented with six word forms and three definitions and have to match the definitions with the correct word form; hence, the test measures form recognition knowledge. The VLT has undergone several changes (Beglar & Hunt 1999; Nation 1983, 1990; Schmitt, Schmitt & Clapham 2001) and recently an updated version known as the new VLT has been proposed by Webb, Sasao and

Balance (2017). Rather than as an overall estimate of vocabulary size, the VLT score is presented for each level tested according to a pre-defined mastery criterion (see footnote 29 in section 3.1.2) and thus produces a profile of vocabulary knowledge. The VLT was originally designed as a diagnostic test (Beglar & Nation 2014) and works well with learners at lower levels (Schmitt 2010).

In contrast to the VLT, the Vocabulary Size Test (VST, Nation & Beglar 2007) was designed to give an overall estimate of vocabulary size. It uses a multiple choice format with four response options to test meaning recognition knowledge of the 14,000 most frequent English word families based on the BNC frequency lists (Nation 2006). A validation study by Beglar (2010) showed promising results, but a study by Gyllstad, Vilkaité and Schmitt (2015) showed that the VST was prone to guessing. Later, a second version (Nation 2012b) based on Nation's 20K BNC/COCA lists (Nation 2012a) was designed and several bilingual versions (e.g. Elgort 2013; Karami 2012; Le Nguyen & Nation 2011, see also Paul Nation's homepage https://www.victoria.ac.nz/lals/about/staff/paul-nation) are available as well. The VocabLab test (Peters, Velghe & Van Rompaey 2015) uses a multiple choice format similar to the VST but also includes a fifth 'I don't know' response option. The test measures knowledge of the 5,000 most frequent lemmas and has mainly been used in research by Peters and colleagues (e.g. Peters, Heynen & Puimège 2016; Peters 2018).

In contrast to written receptive vocabulary knowledge, knowledge of spoken word forms is tested much less frequently (Daller, Milton & Treffers-Daller 2007: 5, see also section 3.1.2) but by now several measures have been proposed. In an effort to directly compare aural and written vocabulary size, Milton and Hopkins (2005) developed an aural version parallel to X-Lex called Aural Lex. Like X-Lex, A-Lex uses a Yes/No format with aural prompts and is scored in the same way. While Milton and Hopkins (2006) report good reliability estimates for A-Lex, the use of a Yes/No format with aural prompts is questionable in practice.[48] Recently, a listening vocabulary levels test (LVLT, McLean, Kramer & Beglar 2015) has been proposed, which, despite the name, uses a format more similar to the VST than the VLT and thus taps into aural meaning recognition. Test takers hear a word in a short, non-defining sentence and then tick the correct L1 translation among four multiple choice options. First validation evidence by McLean, Kramer and Beglar (2015) appears to be very promising, but Kremmel (2017) argues that it could at least partially be an effect of the method used.

Another test that uses aural prompts is the Peabody Picture Vocabulary test (PPVT-IV, Dunn & Dunn 2007). It differs from the aforementioned examples because it is a picture-based test designed to be used with young children and is mainly employed in L1 research (Schmitt 2010) and clinical settings (Kremmel 2017). For reasons of design and item sampling, Kremmel (2017)

---

[48] A-Lex has been piloted for use in the present study with a group of PhD students and postgraduate researchers in English studies with various L1 backgrounds. The results and the feedback indicate that even these L1 and highly advanced L2 users of English found the format difficult and were not able to distinguish between target words and pseudowords solely based on aural prompts, although they would have had knowledge of the target words. While this may be a result of the artificially sounding RP pronunciation used in the A-Lex prompts, it calls the whole test format into question.

argues that the PPVT is of limited value in L2 vocabulary research; however, the idea of using pictures to test the vocabulary knowledge of young pre-literate learners has been taken up by L2 researchers and more recently a further picture-based measure using audio prompts, the Picture Vocabulary Size Test (PVST), has been proposed by Anthony and Nation (2017).

As can be seen from Table 3.7, few productive vocabulary size tests are available in comparison to tests of receptive vocabulary size. In addition to translation tests and free production measures, only two tests, the Productive Vocabulary Levels Test (PVLT, Laufer & Nation 1999) and Lex30 (Meara & Fitzpatrick 2000), have been proposed. Free production measures, such as the Lexical Frequency Profile (LFP, Laufer & Nation 1995) or computational indices of lexical diversity and sophistication (e.g. Crossley et al. 2011; Jarvis 2013) have been characterized as "indirect vocabulary assessments, in which information about a person's lexical knowledge is gleaned from pieces of speech or writing" (Kremmel 2017: 33). They do not provide concrete size estimates and are therefore not discussed here.[49]

The PVLT (Laufer & Nation 1999) is based on the VLT (Nation 1990) and tests the same word frequency levels using a gap-fill task. Test takers are presented "with a meaningful sentence context and with the first few letters of each target word" (Beglar & Nation 2014: 176), although the number of letters given varies (Schmitt 2010). The PVLT thus measures cued form recall, which is why Laufer and Nation (1999: 36) originally called it a test of "controlled productive ability". Although the PVLT has been used relatively frequently in research, concerns have been raised about the construct of the test. Read (2000: 126), for instance, argues that "the blank-filling version may simply be an alternative way of assessing receptive knowledge rather than a measure of productive ability". In addition, Schmitt (2010) criticizes a number of design features such as the number of initial letters given or the strength of the collocations included in the test.

A second productive measure is Lex30 (Meara & Fitzpatrick 2000), which is located in between a cued recall and a free production measure. Lex30 uses a word association task: test takers are presented with a set of thirty highly frequent cue words and are asked to produce four words in response, resulting in a maximum of 120 words. The response words are then scored according to a frequency list (see section 5.3.5.4) based on the assumption "that beginning learners will not generally produce low frequency responses in this task, and that the presence of low frequency words in a test-taker's response set indicates that they have an extended productive vocabulary" (Meara 2009: 132). Some validation evidence has been provided for the test and it has been used in a number of empirical studies (see section 5.3.3.3). Lex30 has the advantage of being easy to administer and quick to complete (Beglar & Nation 2014) and requiring much less receptive knowledge than, for instance, the PVLT (Kremmel 2017). However, the test also has two major disadvantages: first, the construct is not entirely clear (Fitzpatrick 2007; Kremmel 2017) and second, the scores do not provide a direct indication of vocabulary size: although they can be used to rank participants' in terms of size, the scores do not provide an overall estimate of how

---

[49] For a basic overview, see, for instance, Schmitt (2010: 205–216).

many words a test taker knows productively (Walters 2012). At present, the lack of score interpretability is the greatest drawback of using Lex30 and limits its practical usefulness for many pedagogical and research purposes; however, Meara and Olmos Acoy (2010) have proposed innovative ideas based on a capture-recapture methodology used in ecological research which could circumvent this problem and will hopefully be explored further in the future.

Finally, one test that measures vocabulary size and depth in an integrated way has been included as well. Laufer and Goldstein's (2004) Computer Adaptive Test of Size and Strength (CATSS) uses four different test formats corresponding to meaning recognition, form recognition, meaning recall and form recall. The test includes a total of 600 items at four different frequency levels and an academic vocabulary level, but adapts the items presented to test takers in two ways: first, test takers are presented with the most difficult level of knowledge of the form-meaning link (form recall) and if that is known lower levels of knowledge are not tested. Second, the test progresses relatively quickly through the frequency levels until test takers make mistakes, then this frequency level is tested more extensively to provide a more detailed picture of their vocabulary size. This twofold adaptiveness presents a major advantage of the CATSS as it avoids wasting test takers' time (Schmitt 2010).

Albeit brief, this section highlighted that the difficulty of measuring vocabulary is not to be underestimated. Construct definition crucially depends on individual researchers' answers to the underlying question of what it means to know a word, as discussed in section 3.1.1, and the test purpose they envisage. The process of test design entails many decisions about sampling of target items, test format, scoring and the presentation of results, which each include different pitfalls. Moreover, it is essential that test designers then provide comprehensive evidence for the claims of validity, reliability and usefulness made, which is currently done far too little, to repeat Kremmel's (2017) critique. Not surprisingly, currently available size tests display great variability in relation to construct definition, design aspects and the amount of validation evidence available. The broad overview given in section 3.2.3 shows that most size tests measure receptive knowledge of written word forms and the form-meaning link, while fewer tests have investigated aural receptive vocabulary size or written productive size.

3.3 Previous research on extramural English and vocabulary learning

Having discussed L2 lexical development with a focus on vocabulary size and issues of vocabulary measurement, this section provides a more detailed review of previous studies on vocabulary acquisition from extramural English (section 3.3.1) as well as vocabulary uptake from specific activities such as reading (section 3.3.2.1), listening (section 3.3.2.2), viewing (section 3.3.2.3) or gaming (section 3.3.2.4), in which learners frequently engage in informal contexts.

3.3.1 Vocabulary learning from extramural English

Studies on the relationship between extramural English and vocabulary have emerged from two directions: researchers interested in informal language learning who also explore vocabulary knowledge (Lee & Dressman 2018; Olsson & Sylvén 2015; Sundqvist 2009a) and researchers focusing on vocabulary who became interested in out-of-school learning opportunities (Milton 2008; González-Fernández & Schmitt 2015; Peters 2018; Schmitt & Redwood 2011). Studies belonging to the first group have already been mentioned in Chapter 2 and are taken up again in this section, while most of the latter are introduced here for the first time.

In the field of vocabulary research, two studies by Schmitt and Redwood (2011) and González Fernández and Schmitt (2015) have raised interest in out-of-school language input, although they investigated learners' knowledge of English phrasal verbs and collocations without an explicit focus on EE. Both studies examined several potential influencing factors and used a background questionnaire which included questions concerning everyday engagement with English. Schmitt and Redwood (2011) found that informal contact with English through reading and watching films and television had a significant impact on 68 international students' knowledge of phrasal verbs, whereas variables regarding formal instruction did not. Similarly, González Fernández and Schmitt's (2015) findings show that out-of-class contact with English among 108 Spanish EFL learners correlated more strongly with their knowledge of collocations than the number of years they had studied English. While these studies are not concerned with extramural English per se, their results are certainly of great interest with regard to informal vocabulary acquisition.

Among the studies with an explicit focus on extramural or informal leisure time activities, the project led by Berns, de Bot and Hasebrink (2007) was one of the earliest and one of the largest studies to date. Data from 2,248 secondary school learners of English were collected in Belgium, France, Germany and the Netherlands between 1995 and 2000. Among other measures, participants completed a background questionnaire and a Yes/No test (based on Meara 1992). Results show that already towards the end of the 1990s music, TV, computers and the cinema were popular points of contact with English. Vocabulary size as measured by the Yes/No test was highest in Flanders (76.7/100) followed by the Netherlands (66.8/1000) and Germany (53.5/100)[50] and there was a small, but significant correlation between vocabulary size and entertainment media such as music and films.

This early study was followed by a series of Swedish studies also focusing on secondary school learners. Sylvén (2004/2010)[51] set out to compare the vocabulary knowledge of 363 16- to 17-year-old CLIL and non-CLIL students in a longitudinal study spanning over two school years. She administered a questionnaire and a variety of vocabulary measures (cloze test, multiple-choice test, VKS, and a words in context test in combination with reading activities) and found that while CLIL students generally scored higher than non-CLIL students, students who received

---

[50] Unfortunately, the vocabulary test was not administered to the French participants.
[51] Sylvén's study first became available as a thesis in 2004 but was published as a book in 2010.

English input outside school scored best overall. On average, male students had higher scores, which could also be a mediating effect of out-of-school input because boys reported more frequent contact with English. In addition, Sylvén (2004/2010) found that reading was an especially valuable EE activity, as readers performed better in all groups than non-readers.

Sundqvist (2009a) further explored Sylvén's (2004/2010) unexpected findings with regard to extramural English. Focusing on oral proficiency as well as vocabulary, she gathered data from 80 EFL learners in grade 9 (15 to 16 years) with the help of a questionnaire, structured language diary, the VLT (2K, 3K and 5K level) and a shortened version of the PVLT among other instruments. Her results show that listening to music, playing video games and watching TV and films were the most popular EE activities and that, on average, students spent 18.4 hours with extramural English per week. Regarding vocabulary knowledge, she found significant correlations between an EE index variable and the scores on the receptive (r = .354) and the productive vocabulary test (r = .352). However, further analysis of the vocabulary scores showed that boys performed better than girls on both tests and that the correlations between the vocabulary measures and the EE index were only significant for boys, but not for girls, if the sample was split according to gender. In a backward regression analysis, Sundqvist (2009a) found that playing games and surfing the internet contribute most to vocabulary learning, followed by watching TV, reading newspapers and magazines, reading books, listening to music and watching films.

Olsson (2012) investigated the relationship between extramural English and student writing with two classes of learners in grade 9 (*N* = 37). She collected a corpus of 74 learner texts, which were analysed with regard to vocabulary. EE data were again collected using a questionnaire and language diary and showed similar trends to the previous Swedish studies: listening to music was the most common activity with 86% of participants reporting that they engaged in it every day, 50% watched TV programmes with subtitles on a daily basis, and 41% regularly spoke English outside school, whereas there was hardly any writing outside school. In general, boys reported more EE activities than girls; the clearest gender difference was found for gaming, which was done much more frequently by boys, while reading was more popular among girls. Overall, participants' mean EE time ranged between 18 minutes and more than 7 hours per day with an average of 2.9 hours per day, which amounts to 20.3 hours a week. With regard to the use of vocabulary in learner texts, Olsson (2012) found that students with more exposure to English presented a more varied (informal) vocabulary (operationalized as type/token ration) in a letter writing task and tended to use longer words and more infrequent words (> 3K frequency level) in articles they wrote. Hence, overall her results suggest that learners who engage more with English outside school have a richer vocabulary.

The relationship between EE and vocabulary in learner texts was explored further in a larger study by Olsson and Sylvén (2015), which focused on the development of academic vocabulary among CLIL and non-CLIL students over a period of three years. In total 525 essays by 230 students in upper secondary school (16 – 19 years) were analysed. CLIL and non-CLIL students

were also compared in terms of their engagement with EE: data from a survey and language diary show that the CLIL students (*n* = 146) used significantly more English outside school than non-CLIL students (*n* = 84), especially with regard to reading and writing. In addition, boys had significantly more contact with English than girls; however, on closer inspection this difference is explained by only one activity – gaming – while there is no gender difference for other activities. Participants in Olsson and Sylvén's (2015) study reported spending an extraordinary amount of time with English: on average, CLIL students were in contact with English for more than 7.5 hours a day and non-CLIL students for more than 5.5 hours.[52] In terms of academic vocabulary, results show that male CLIL students used the highest proportion in their texts, whereas female non-CLIL students used the least. However, regression analyses controlling for baseline differences at test time 1 show that all groups developed similarly in their ability to use academic vocabulary. Thus, while there was a significant correlation between academic vocabulary use and frequency of EE engagement in the first assignment, these initial group differences did not increase over time and none of the later three assignments showed significant correlations with EE.

Taken together, these four Swedish studies (Olsson 2012; Olsson & Sylvén 2015; Sundqvist 2009a; Sylvén 2004/2010) show some trends with regard to extramural English in general and its relation to vocabulary in particular: first, Swedish secondary school students spend a substantial amount of their leisure time in contact with English, which even seems to have increased over the years. Language diary data collected in 2006 and 2007 shows a mean EE time of 18.4 hours (Sundqvist 2009a), in data from 2009 it had increased to 20.3 hours a week (Olsson 2012) and in Olsson and Sylvén's study (2015) even non-CLIL students reported a total of 39.2 hours a week.[53] Second, the most common EE activities in these studies seem to be rather stable and revolve around popular media: music is generally most frequently engaged in, followed by TV/films and the internet. These activities clearly favour receptive skills, and it appears that few students regularly speak or write English outside formal educational contexts. Furthermore, the studies have found large gender differences concerning gaming, which also seems to be the only popular activity that uses language in a productive way. Third, with regard to vocabulary the four studies suggest that EE has an impact on vocabulary knowledge: it correlates with productive and receptive vocabulary size (Sundqvist 2009a), the use of a more varied vocabulary and more infrequent words in learner texts (Olsson 2012), and to some extent with the use of academic vocabulary (Olsson & Sylvén 2015). However, the results on which activities are most beneficial for lexical development are somewhat contradictory: while Sylvén (2004/2010) suggests that reading might play a fundamental role, Sundqvist (2009a) found that playing digital games and surfing the internet were the best predictors of vocabulary knowledge.

---

[52] The difference in mean EE time was significant for CLIL and non-CLIL students, whereas there was no significant difference between male and female participants.

[53] Studies by Sylvén and Sundqvist (2012a) and Sundvist and Sylven (2014), which focus more specifically on gaming (see section 3.3.2.4), include data on time spent with EE among younger Swedish learners and show that already 10- to 11-year-old pupils spend an average of 7.2 hours per week with English outside school.

Research involving secondary school learners has however not only been conducted in Sweden, but also in the Netherlands, Austria, and Belgium. Similar to Olsson and Sylvén (2015), Verspoor, de Bot and van Rein (2011) compared learners in monolingual Dutch and bilingual Dutch-English streams in two school types: denominational schools of Dutch Reformed Churches, whose students have limited exposure to popular media and thus to English outside school, and regular state schools, where no such restrictions apply. In total, 240 students in the first year of secondary school (13 years) and 316 students in the third year (15 years) took part in this study. A questionnaire including items on out-of-school English input and a Yes/No vocabulary test (Meara & Buxton 1987) were administered as part of a larger study, which also measured reading and writing skills. The learners received the same test three times in the course of one academic year: first-year students were tested on their knowledge of the 2,000 most frequent words, while third-year students were tested on the 3K and 4K bands the first two times and on the 3K to 5K bands at the third test time. For the first-year students there was no significant difference in participants' vocabulary knowledge at test time 1, but at test times 2 and 3 students in bilingual streams outperformed the students in the monolingual streams. However, the students in bilingual streams with exposure to English outside school also performed significantly better than their counterparts in the bilingual streams of religious schools. For the third-year students, differences in vocabulary were found at all test times: again, the students in bilingual streams performed better at all three test times, but students in monolingual streams with access to English media always scored significantly higher than the monolingual group whose access to media was limited. For the bilingual media/non-media groups a significant difference was only found at test time 3, again in favour of the students with access to English media outside school. Correlational analyses showed no significant relationship between popular media and lexical knowledge for the first-year students at test times 1 and 2 and a relatively weak correlation ($r = .225$) at test time 3, but only for bilingual students. For the third-year students, significant correlations were found at almost all test times with a maximum strength of $r = .305$ for the monolingual group (time 2) and $r = .478$ for the bilingual group (time 3). Drawing on the results for lexical proficiency as well as reading and writing, Verspoor, de Bot and van Rein (2011: 163–164) conclude that "[o]ur main finding is that the lack of input from the media had a long term effect on the developing proficiencies of the non-media group as English language learners."

Another project on secondary school students was conducted in Austria by Hahn (2017, 2018). This study deserves special attention for two reasons: first, it is one of few projects that focuses on students in vocational middle schools rather than students in academic schools studying for their A-levels;[54] and secondly, it is an approximate replication (Porte 2012) of the quantitative strand of the present study. Hahn (2017) also focused on students in grade 10 and collected data using the same quantitative instruments (see section 5.3.3) with the exception of the language diary, which could not be administered for practical reasons. Therefore, the two studies are

---

[54] See section 4.2.1 for more information on the Austrian school system.

directly comparable despite their different research contexts. A sample of 83 10th-grade students in three Viennese vocational business middle schools (HAS) between the ages of 15 and 18 participated in Hahn's (2017) study. The students, who are supposed to have reached level A2 of the CEFR (Bundesministerium für Bildung 2014), had a mean receptive vocabulary size of 3,041 words as measured by V_YesNo, and achieved an average score of 23.6 on the Lex30 test.

Results on the types of EE activities show that music is the most important point of contact for these learners: approximately 70% listen to music on their phone or mp3 player at least a few times a week, watch music videos and sing along to music, and almost 50% use streaming services for music. In addition, more than 40% also read the lyrics or translations. The second most frequent source of English input are audiovisual media with 70% watching online video clips, over 60% watching films and over 50% watching series online at least a few times a week. In addition, reading English on social media constitutes a third type of contact with more than 50% each reading status updates and comments or messages at least a few times a week (Hahn 2017: 78–79). EE exposure correlated significantly with receptive (r_s = .398, p = 0.001, N = 67) and productive (r_s = .341, p = 0.002, N = 83) vocabulary size and significant differences were found between high and low EE exposure groups for both vocabulary measures. Similar to the Swedish studies, boys engaged in significantly more EE activities than girls both on a weekly and daily basis, and significant differences were found for gaming and use of social media. Moreover, boys also had a significantly larger receptive vocabulary, whereas there was no statistically significant difference for Lex30 scores. Overall, Hahn found similar trends to the previous studies in Sweden with regard to EE, vocabulary knowledge and gender differences, which indicates that EE exposure in non-subtitling countries such as Austria is not entirely different from subtitling countries such as Sweden and that students in vocational middle schools show similar behaviour to students in academically-oriented schools.

Peters (2018) compares vocabulary knowledge and out-of-class exposure to English among 47 16-year-old students in their fourth year of secondary school and 32 19-year-old students in their first year at university in Flanders, Belgium. She used a questionnaire and the VocabLab test (Peters, Velghe & Van Rompaey 2015, see section 3.2.3). Similar to the studies discussed above, Peters (2018) found that the most commonly used sources of extramural English were songs, subtitled movies, subtitled TV programs and the internet. Films and TV programs without subtitles were more popular among the older participant group, whereas the 16-year-olds engaged more frequently in gaming. In terms of vocabulary knowledge, test scores indicate that the 16-year-old learners had mastered the 2K frequency level and the 19-year-old university students both the 2K and the 3K levels. This is remarkable because at the time of study, these students had only been studying English in formal education for three and six years respectively.[55] However, exposure to English outside educational contexts in Flanders is generally quite high (see also De Wilde & Eyckmans 2017), which is reflected in the results of

---

[55] Because French is the first foreign language taught in Flanders, English is usually only taught from the beginning of secondary school (De Wilde & Eyckmans 2017; Peters 2018).

this study: positive, though relatively small correlations were found for movies and TV programs without subtitles, reading books and magazines and using English-language websites. A composite exposure index correlated with vocabulary size at r = .35 and explained 13% of variance in vocabulary scores in an ANCOVA analysis, whereas length of instruction only explained 7%. In addition, there was no effect for gender. Thus, Peter's (2018) findings extend those of previous studies by showing that exposure to English in informal contexts had more predictive power in relation to vocabulary size than length of instruction. Clearly, this could be a result of the specific research context, but it is worth remembering that González Fernández and Schmitt (2015) found a similar result with Spanish learners in their study on phrasal verbs.

An even more recent study by Peters et al. (2019) also included Flemish learners at secondary school and university level, but compared their knowledge of French and English as a foreign language and investigated the influence of length of instruction and gender in addition to EE exposure. 138 participants in the second year of secondary education (n = 48), the fourth year of secondary education (n = 43) and the first year of university (n = 47) took part in the study. Their receptive vocabulary knowledge was measured using an English and a French version of the VocabLab test and EE data was collected using a questionnaire. Results show that participants have significantly more exposure to English than to French through media. A significant gender difference was found for English with male participants engaging in more online activities such as gaming and surfing the internet. Furthermore, participants at all three educational levels had significantly larger vocabulary sizes for English than for French, despite the fact that formal instruction in French begins three years earlier than instruction in English. In a structural equation model for English test scores, educational level, which relates to length of instruction, was found to be a significant predictor of vocabulary knowledge, whereas of six EE-related factors only the factor of gaming and online activities had a significant effect. Gender had an effect on this EE gaming/online factor, but not directly on vocabulary size. This study thus shows that despite a longer period of instruction for French as the first foreign language, Flemish learners have greater vocabulary knowledge in English than in French due to the amount of EE exposure. In fact, the authors explain the lack of significance for all out-of-school factors other than gaming and online activities in the following way:

> Yet, it would be incorrect to conclude that the other variables are not beneficial for learners' vocabulary growth. The fact that no other variables had an impact on participants' vocabulary knowledge could be explained by participants' large amounts of out-of-school exposure to English and, consequently, by a lack of variance because of a ceiling effect for a number of questionnaire items. For instance, only one participant reported never watching subtitled TV programs; three participants reported never watching subtitled movies (Peters et al. 2019: 26).

Further studies on learners at university level were conducted by Cole and Vanderplank (2016) and Lee (2019a, see also Lee & Dressman 2018).[56] The first study explores the linguistic knowledge of 34 Fully Autonomous Self-Instructed Learners (FASILs) in comparison to 50

---

[56] Kaur's (2015) study (see section 2.3) also included university level learners, but since her study explored learners' use of opportunities for vocabulary learning in out-of-class contexts in a qualitative study and does not include data on vocabulary uptake, it is not included in this review.

Classroom-Trained Learners (CTLs) in Brazil. The existence of FASILs, who generally begin learning English "as by-product of committed engagement with informal sources of English such as television and music" and then continue "using English in their spare time and participat[ing] in activities in which English [is] the shared language, especially in online environments" (Cole & Vanderplank 2016: 34), was discovered in an earlier exploratory study. FASILs can be described as EE-only or mainly-EE learners and Cole and Vanderplank (2016) aim to describe their competences and possible advantages over CTLs using a variety of measures. Among these, vocabulary use was assessed as the lexical resources available for the oral re-telling of a video clip. Results show that FASILs performed significantly better than CTLs with regard to vocabulary use in their oral narrative as well as several other skills. FASILs also spent more time with informal English: over 50% of the participating FASILs used English for more than two hours a day compared to around one hour per day or less for the majority of the CTLs. However, a multiple regression analysis showed that the amount of time spent with informal English did not have a significant effect on proficiency in this study.

Lee's (2019a) study is more directly related to vocabulary acquisition: he explored the relation between informal digital learning of English (IDLE) activities (see section 2.2) and vocabulary knowledge. Lee (2019a) administered a questionnaire and the University Word List (UWL) section of the VLT and PVLT to 94 Korean learners of English (mean age = 22) and conducted interviews with 77 participants. Data from the questionnaire show that almost 50% of the participants spent less than one hour with informal English activities a day and slightly more than 25% one to two hours per day. The number and quality of these IDLE activities was assessed on the basis of the semi-structured interviews. When correlating the scores on the UWL section with information on participants' IDLE activities, Lee (2019a) found that while the scores did not correlate significantly with the quantity of IDLE activities, both the receptive (r = .27) and productive (r = .46) vocabulary scores did correlate with quality. His results therefore suggest that, at least for Korean learners of English, a mix of form- and meaning-focused is more important than only the number of the activities overall.

Research on EE and vocabulary learning has not only been conducted with learners at secondary school or university, but also with young language learners in primary school or children who had not yet received any formal instruction. Several studies have been conducted with primary school pupils in Flanders, Belgium, where the first foreign language taught is French and pupils do not receive any formal English instruction at primary level. However, "Flemish children [...] are often exposed to English outside the school from an early age onwards" (De Wilde & Eyckmans 2017: 675), mainly due to subtitled English-language TV broadcasts. Consequently, any English that these children know has likely been learned from extramural English, which makes these studies particularly interesting. An early example is the study by Kuppens (2010), who investigated knowledge of vocabulary in two translation tests with 473 Flemish pupils with L1 Dutch in the last year of primary education and found that watching English-language TV

and movies and to a lesser extent playing computer games significantly contributed to students' scores.

More recently, De Wilde and Eyckmans (2017) carried out a pilot study with 30 11-year-old pupils in the last year of primary school and De Wilde, Brysbaert and Eyckmans (2019) conducted a further large-scale project with 780 pupils, who were also in their last year of primary education. In both studies, the researchers used the Peabody Picture Vocabulary Test (PPVT-4, Dunn & Dunn 2007) to explore children's receptive English vocabulary knowledge and also measured the skills of listening, reading/writing and speaking. Data from questionnaires given to the children and their parents in the first study (De Wilde & Eyckmans 2017) show that 60% of the children listened to English-language music for more than one hour a day, 40% of the children used English on the computer for more than one hour, over 30% spent more than an hour gaming, and nearly 60% watched TV programmes with subtitles for 30 minutes to an hour or more on a daily basis. In the second large-scale study (De Wilde, Brysbaert & Eyckmans 2019), 97% of the children listened to English-language music every day, 80% watched TV with subtitles, 78% used social media and 75% played digital games on a daily basis, whereas there was hardly any reading of print media. This seems to be a considerable amount of English input for young learners and in fact, the results of the skills-based tests in both studies show that many of the children could already perform English tasks at A2 level and that listening comprehension was particularly good, although there was a wide range of scores on all measures. More specifically, De Wilde, Brysbaert and Eyckmans (2019) found that 25% of the pupils attained A2 level for listening comprehension, 14% for the speaking test and 10% for the integrated reading and writing test.

Regarding vocabulary a similar picture was found, in the pilot study (De Wilde & Eyckmans 2017) the mean score on the PPVT-4 was 66.2 out of 108, but scores ranged from 12 to 102 within the relatively small sample. In the second study (De Wilde, Brysbaert & Eyckmans 2019) the mean score was 78/120 with a minimum of 31 and a maximum score of 116. Concerning influencing factors on vocabulary knowledge, no significant effects were found for gender, parental education or home languages in the pilot study (De Wilde & Eyckmans 2017), but there was an effect for two EE activities: gaming and computer use were significantly related to the vocabulary test scores. For the large-scale study (De Wilde, Brysbaert & Eyckmans 2019), a multiple regression analysis showed that using social media, playing English games and speaking English had a positive and significant effect on the children's receptive vocabulary knowledge, but there was also a small negative and significant effect of listening to music.[57] Overall, EE activities explained 18% of the variance in the vocabulary scores. Surprisingly, no effect was found for subtitled television in either study, although it constitutes a major source of English input.

---

[57] For the results of the regression analyses in relation to the skills-based tests as well as an overall index of language proficiency, please see the article by De Wilde, Brysbaert and Eyckmans (2019).

Puimège and Peters (2019) also carried out a large-scale study with 616 Flemish primary school pupils between the ages of 10 and 12, who also had not received any formal English instruction. They investigated both learner-related and word-related factors using a questionnaire to collect EE data, a Dutch vocabulary test, and the Picture Vocabulary Size Test (PVST) by Anthony and Nation (2017) to test meaning recognition in English. The target words used in the PVST were also administered in a format requiring translation or explanation to test meaning recall. Results of the English vocabulary measures show that both meaning recognition and recall knowledge increase with age from a mean of 2,356 word families among the 10-year-olds to a mean of 3,157 word families among the 12-year-olds for meaning recognition and from a mean of 1,295 to 2,125 word families for meaning recall. Similarly, engagement to EE, which was summarized in the three factors passive exposure, gaming and streaming, and reading and writing according to a principal component analysis, increases with age as well. Concerning factors affecting vocabulary knowledge, cognateness was found to be the most influential word-related factor, but frequency also showed significant positive correlations with both meaning recognition and recall, whereas concreteness only predicted meaning recognition and part of speech showed no significant effect for either test. Among the learner factors included in the study, age and L1 (Dutch) vocabulary knowledge were significant predictors for both meaning recognition and recall. Gender also affected both tests with boys performing better than girls. In addition, meaning recognition was positively related to learner's passive EE exposure, i.e. listening to music and watching subtitled TV, whereas the EE factor gaming and video streaming was a positive predictor of meaning recall.

Overall, these studies suggest that Flemish children are able to acquire considerable competence in English based on EE input, although large variations were found. The latest study by Puimège and Peters (2019) presents an important extension of previous research because it is one of the first studies on EE and vocabulary which systematically explored word- and learner-related factors.

Persson and Prins (2012) conducted a study in a comparable environment in the Netherlands and investigated the vocabulary knowledge of even younger primary school pupils with and without early English instruction. 178 early English learners and 33 non-instructed learners of English with a mean age of 4.5 years took part in the project. Among other tests, their English vocabulary size was measured at the beginning and end of the school year using the PPVT-4. Information on out-of-school English exposure was collected by means of a parental questionnaire. Results show that watching English TV and playing digital games were the most common activities, but exposure varied greatly. Concerning vocabulary size, there was no significant difference between the early English and the no English group at test time 1, but the early English group improved significantly over the course of the school year, while the learning gains of the no-English group were not significant. A regression model for the early English group showed that children's age at the time of testing, quality of English input at school and exposure to TV programmes produced for English-speaking children were significant predictors

of vocabulary size, whereas quantity of input at school was not significant. Referring to the non-instructed learners of English, Persson and Prins (2012: 10) conclude that "even in minimal input situations the English vocabulary and grammar of young Dutch foreign language learners improves in one school year."

Finally, Jóhannsdóttir (2017) investigated the vocabulary knowledge of 378 fourth-grade primary school pupils (8 to 9 years) in Iceland, who received minimal English instruction of 40 minutes per week in grades 1 to 4. Knowledge of 100 target words was tested using a Yes/No test and 25 of these target items were also administered in the VKS format (Paribakht & Wesche 1993, 1997). In addition, all pupils filled in a questionnaire focusing on motivation, which also included questions on extramural exposure to English. Not surprisingly for this age group, the most popular activities were listening to music and playing computer games. L2 learning experience correlated significantly with the scores of the Yes/No test (r = .427) and the VKS (r = .444); in fact, all EE activities included in the study (listening to music and TV, using computers, reading texts and using English as a lingua franca) showed significant correlations with both test scores, whereas using English at school did not. Jóhannsdóttir (2017: 72) thus argues that her results suggest "a primarily media exposed youth acquiring English incidentally through recreational activities rather than focused learning such as at school".

The studies reviewed in this section show that research investigating the relationship between extramural English and vocabulary has tended to concentrate on vocabulary size. Most of the studies use cross-sectional research designs with one test time and a focus on correlations, only four studies are longitudinal in nature and measured vocabulary at at least two points in time (Persson & Prins 2012; Olsson & Sylvén 2015; Sylvén 2004/2010; Verspoor, De Bot & Van Rein 2011). Moreover, vocabulary measurement could be improved in some studies in which the construct of word knowledge is not entirely clear and tests have been adapted without subsequent validation (Jóhannsdóttir 2017; Sundqvist 2009a) or very small samples of target items have been used (Lee 2019a). While such measurement issues are certainly also found in other studies, they reflect the fact that many of these researchers are not primarily concerned with vocabulary, but with informal language learning.[58]

Regarding the findings, similar EE activities were found to be popular in the different samples, especially among teenagers and university students. Although the exact order of popularity varies, exposure to English outside formal educational contexts clearly revolves around media such as music, TV programmes and films, and games and often involves a screen. Data from Sweden show that teenagers without special English education, such as CLIL, spend between 2.6 (Sundqvist 2009a) and 5.6 hours a day (Olsson & Sylvén 2015) with English and that the amount of EE appears to have risen over time.

Studies on young language learners without formal English instruction have established that children can acquire English vocabulary solely based on the input they receive outside school

---

[58] See also the comments about test usefulness for users of vocabulary tests in section 3.2.3.

(De Wilde & Eyckmans 2017; De Wilde, Brysbaert & Eyckmans 2019; Persson & Prins 2012; Puimège & Peters 2019). However, English teaching at school significantly improved vocabulary knowledge in Persson and Prins' (2012) study and quality, but not quantity, of input at school was a significant predictor of vocabulary size; hence, it is important not to overstate claims of naturalistic learning in out-of-school contexts. In addition, Cole and Vanderplank (2016) suggest that naturalistic, informal language learning in online contexts is not only possible for children, but also for adults using the example of their FASILs.

At secondary school level, data from Sweden (Olsson 2012; Olsson & Sylvén 2015; Sundqvist 2009a; Sylvén 2004/2010), Flanders (Peters 2018; Peters et al. 2019) and the Netherlands (Verspoor, De Bot & Van Rein 2011) show that exposure to English outside school benefits vocabulary development, which is perhaps little surprising as these countries are known for their early exposure to English, particularly through English-language TV broadcasts with L1 subtitles. However, the evidence regarding effects of subtitled TV is inconclusive, with positive effects in some studies (Puimège & Peters 2019) but not in others (De Wilde & Eyckmans 2017; De Wilde, Brysbaert & Eyckmans 2019; Peters et al. 2019). In addition, similar results have also been found in Austria (Hahn 2017), which is a dubbing country. These findings suggest that subtitled television is not the decisive source of English input any more, particularly since learners have a wide range of online resources at their disposal.

Investigating informal learning activities at university level, Lee (2019a) found that rather than the quantity of these activities, it is their quality which is related to vocabulary knowledge, echoing the results of Persson and Prins (2012). Furthermore, in a comparative study of learners at upper secondary school and university Peters (2018) showed that length of formal instruction is not as good a predictor of vocabulary size as exposure to extramural English, which is a highly interesting finding supported by previous studies (González-Fernández & Schmitt 2015; Schmitt & Redwood 2011). Peters et al. (2019) further corroborate this result by showing that despite a longer period of instruction for French, Flemish learners of French and English have significantly larger English vocabulary sizes.

In sum, this synthesis of currently available research suggests that engagement with English in informal, out-of-school contexts has an impact on vocabulary development which may even be more profound than the effects of formal English instruction. However, more research needs to be done to establish the scope of this impact in relation to different word knowledge aspects and in different learner populations.

3.3.2 Vocabulary learning from specific activities

Rather than in relation to students' overall out-of-school English input, vocabulary learning can also be researched in relation to specific activities. Although most of the studies presented in this section do not use the terms extramural English or informal learning, they are highly informative as they investigate activities that learners engage in during their leisure time, such as listening, viewing, reading or gaming. Most of the studies reviewed use the term incidental

vocabulary learning (see section 3.1.2) to describe their object of investigation, but some researchers use related terms such as contextual word learning (Elgort et al. 2018) or lexical inferencing (Van Zeeland 2014), when a differentiation according to intentionality is not of particular interest. An effort has been made to incorporate all relevant studies in this review, but since the literature on vocabulary acquisition is vast, there will be some necessary gaps. For instance, studies conducted in educational settings have only been included if they had a clearly discernible focus on incidental learning, whereas studies on intentional learning or instructed vocabulary learning are not considered at all.

3.3.2.1 Vocabulary learning from reading

Most research in the area of incidental vocabulary learning has been conducted in relation to reading, which is why it is virtually impossible to provide an exhaustive overview of all previous studies. Therefore, this review focuses on studies using longer texts and research on extensive reading, which some learners may engage in outside school, as well as on more recent studies because they present interesting methodological advances, such as the use of eye-tracking (Elgort et al. 2018; Mohamed 2018; Pellicer-Sánchez 2016) or the testing of implicit lexical knowledge (Elgort & Warren 2014).

In general, earlier studies on vocabulary learning from reading have shown that learners can pick up words incidentally, but learning gains are relatively small (Waring & Nation 2004). Drawing on a summary of learning rates in empirical studies by Waring and Takaki (2003), Waring and Nation (2004) note that gains vary greatly between 5% and 25%, which they attribute to the target words tested, the method of measurement and the type of word knowledge tested. They thus advocate using multiple measures to obtain a more detailed picture of the development of different types of word knowledge from reading, which is taken up in most of the more recent studies.

Waring and Takaki (2003) conducted a study using a graded reader with 15 Japanese university students. 25 target words in five frequency bands were tested using three measures of form recognition, meaning recognition and meaning recall in an immediate and two delayed post-tests. Mean scores on the immediate post-test show that form was recognized for 61.2% of the target words, while meaning could be recognized for 42.4% and recalled for 18.4%. 54.9% of the words known at form recognition level were still retained three months later and 57.5% of the words known at meaning recognition level, but only 19.5% of the words known at meaning recall level. While this result is influenced by the low learning rate at recall level overall, Waring and Takaki (2003) also found an influence of frequency of occurrence of the target items in the text: words that were encountered more than eight times had a 50% chance of being recognized with regard to form and/or meaning three months later; however, even words that were met more than 18 times only had a 10-15% chance of being known at meaning recall level on the second delayed post-test, which suggests that even with repeated exposure meaning recall knowledge is very difficult to acquire from reading.

Pigada and Schmitt (2006) traced vocabulary learning from graded readers over one month in a single-learner case study. A Greek learner of French read four graded readers with a total of 113 French target words and was tested on form recall, meaning recall and grammatical knowledge in a pre- and post-test interview. Results indicate that substantial learning occurred as "knowledge of 65% of the target words was enhanced in some way" (Pigada & Schmitt 2006: 1). More specifically, between the pre- and post-test, spelling, hence form recall, improved by 23%, meaning recall by 15.4% and grammatical knowledge by 30% for nouns and 16.6% for verbs; these differences were significant in all cases. The effect of frequency of occurrence is not completely clear for form recall but seems to be linearly related to meaning recall and grammatical knowledge. For meaning, there was "a discernible rise in the learning rate" (Pigada & Schmitt 2006: 18) after ten exposures or more.

Pellicer-Sánchez and Schmitt's (2010) study on incidental vocabulary learning from reading an authentic novel uses an innovative research design: the English-language novel contains words from the Nigerian language Ibo, 34 of which were selected as target items. 20 Spanish university students, who were advanced learners of English but had no knowledge of Ibo, took part in the study. Their acquisition of the Ibo target words was tested in individual interviews with regard to form recognition, meaning recognition, meaning recall and recall of word class. The findings show that these advanced learners could improve their lexical knowledge on all four measures by reading an authentic novel. The acquisition profile illustrates that the largest gains were made for meaning recognition (43%), followed by form recognition (34%), recall of word class (20%) and meaning recall (14%). Not surprisingly, knowledge at the level of recall was more difficult to acquire, but there was a significant effect of frequency of occurrence for all measures. "After more than 10 exposures, the meaning and spelling could be recognized for 84% and 76% of the words respectively, while the meaning and word class could be recalled for 55% and 63%" (Pellicer-Sánchez & Schmitt 2010: 31).

Elgort and Warren (2014) also introduced an innovation in this line of research, as they tested both the implicit and explicit knowledge of 48 pseudowords inserted in five chapters of a book on economy among 48 international students in pre-university English courses in New Zealand. They found modest gains on the meaning recall test used as the explicit measure: a mean of 10 out of 48 target items (20.8%) were answered correctly. However, no stable effects were found for implicit knowledge on two timed lexical decision tasks using form priming and semantic priming. The authors take this results as an indication "that contextual word learning is a slow and incremental process" (Elgort & Warren 2014: 394)

Finally, three studies of interest employed eye-tracking measures to explore incidental vocabulary learning from reading.[59] Pellicer-Sánchez (2016) compared eye movements of L1 and L2 speakers of English while reading a short story containing six pseudowords, which each occurred eight times. The participants were 23 postgraduate students or postdoctoral

---

[59] Since an introduction to eye-tracking methods in vocabulary research is clearly beyond the scope of this thesis, interested readers are referred to Pellicer-Sánchez and Siyanova-Chanturia (2018).

researchers from different L1 backgrounds and 25 British undergraduate students. In addition to the eye-tracking data, participants also completed measures of form and meaning recognition and meaning recall. Pellicer-Sánchez (2016) found that initial reading times were longer for the pseudowords than for control items among both L2 and L1 speakers, but decreased significantly after eight encounters. On the immediate post-test, L2 participants knew 85.5% of the six target items at the level of form recognition, 78.3% at the level of meaning recognition and 60.9% at the level of meaning recall; whereas L1 participants achieved 91.3%, 86.6% and 65.3% respectively. Statistical tests showed that there was no significant difference between L1 and L2 participants with regard to their post-test scores. In addition, a delayed post-test completed by L2 participants two weeks later showed almost the same results (85.7% for form recognition, 75% for meaning recognition, 54.8% for meaning recall). Comparison of the online eye-movement measures and the offline vocabulary tests showed that both L2 and L1 readers had better recall knowledge of words with longer total reading times. A final interesting aspect of Pellicer-Sánchez's (2016) study is that she also asked participants to rate their certainty in relation to their responses on the post-test. Results indicate that both L2 and L1 users were most certain about their responses with regard to meaning recognition followed by form recognition and meaning recall. Pellicer-Sánchez (2016: 122) thus concludes "that receptive aspects of vocabulary are not only easier to acquire but may also lead to higher degrees of certainty".

Elgort et al. (2018) conducted an eye-movement study with a slightly less advanced sample of L2 users of English: 40 Dutch university students who had an English vocabulary size of 8,000 word families or more as measured by the VST read an expository text of 12,000 words which included 14 low-frequency target items. In addition, they completed a sentence reading post-test and a meaning recall task. Again, the eye movement data show that reading times decreased across the first eight encounters with target items, which suggests that orthographic information is acquired relatively fast, but "the eye-movement measures capturing word meaning and integration in the ongoing text indicate clear differences between the target and the control words up to the very end of text, sometimes after 40 readings of the words" (Elgort et al. 2018: 360). This is in line with results of the meaning recall post-test, on which 34% of responses were correct on average. The authors thus suggest that the 60% found by Pellicer-Sánchez (2016) may be too optimistic an estimate for more average learners.

In a third study by Mohamed (2018), 42 international students with diverse L1 backgrounds and a mean vocabulary size of almost 4,000 words as measured on a Yes/No Test (Meara 1992) read a graded reader containing 20 pseudowords and 20 known control words. An immediate post-test incorporated measures of form recognition, meaning recognition and meaning recall. Results based on the eye-movement data show that reading times for the pseudowords were significantly longer than for the control words, but they decreased over several encounters. Mean scores on the post-test indicated that participants had acquired 41.8% of target items at the level of form recognition, 30.3% at the level of meaning recognition, and 12.9% at the level of meaning recall. Combining data from the online and offline measures showed that "[s]ummed

total [reading] times strongly predicted learning success in all vocabulary measures, particularly in form recognition and meaning recall" (Mohamed 2018: 284). In addition, exposure frequency was also a significant predictor of vocabulary scores although it was not as strong as total reading time.

In sum, these studies provide evidence that several types of word knowledge can be acquired from reading. As Waring and Nation (2004) suggested, learning gains clearly depend on the type of knowledge measured with knowledge on immediate post-tests ranging between 30% and 78% for meaning recognition, between 34% and 86% for form recognition, between 13% and 61% for meaning recall and between 23% and 42% for form recall. However, it has to be pointed out that the larger estimates were almost all found in Pellicer-Sánchez's (2016) study, whose participants were very advanced. Particularly for meaning recall all other studies show gains between 13% and 34%, which is considerably lower than her 61%. Furthermore, data from the eye-tracking studies suggest that readers pay more attention to unknown words in reading input, but that these initially long reading times decrease relatively quickly, which indicates that lexical processing during reading changes as a result of frequency of exposure. This is in line with other studies finding that frequency of occurrence is the most prominent influencing factor (see also Chen & Truscott 2010; Webb 2007), although clearly not the only one (see section 3.1.3).

Research on incidental learning from reading has not only been conducted with single words as target items, but also with multiword items. In particular, several studies have researched the incidental acquisition of collocations (Pellicer-Sánchez 2017; Sonbul & Schmitt 2013; Szudarski & Carter 2016; Vilkaitė 2017; Webb, Newton & Chang 2013). Findings indicate that both L1 and L2 users can learn collocations incidentally (Sonbul & Schmitt 2013), that collocations with pseudowords can be learned incidentally at both recognition and recall level (Pellicer-Sánchez 2017) and that nonadjacent collocations, where collocates do not occur directly next to each other in the text, can be learned as well as adjacent collocations at recognition level (Vilkaitė 2017). Since the present study does not include multiword items, further discussion is beyond the scope of this thesis, but detailed information can be found in the studies cited.

A second area of study on vocabulary learning from reading that is highly relevant to the present project is concerned with lexical inferencing (Haastrup 1991), which relates to readers guessing the meaning of unfamiliar words from context (Wesche & Paribakht 2010). Although lexical inferencing does not automatically lead to the acquisition of the lexical items (Elgort 2017), it can contribute to incidental learning, particularly in relation to meaning. Multiple studies have investigated the success of lexical inferencing from written input and Table 3.8 provides an overview of key studies. Since most of the research was conducted with learners of English, all studies are concerned with English as the target language, unless explicitly stated otherwise.

Study	Participants	Text type	Lexical inferencing success		
			Correct	Partially correct	Incorrect
Bengeleil & Paribakht (2004)	Arabic-speaking students of medicine (intermediate to advanced)	Expository text of ca. 1,000 words	41.0%	16.6%	42.4%
Elgort (2017)	Chinese-speaking students (high intermediate)	Single sentences	61.5%		29.2%
Haastrup (2008)	Danish students in 7th and 10th grade, and first year of university (mixed)	Five short factual texts	34.1%	-	-
Nassaji & Hu (2012)	Chinese-speaking students (advanced)	Short introductory academic text	58.1%	-	-
Nassaji (2003)	International students (intermediate)	Short essay	25,6%	18.6%	55.8%
Pulido (2007)	Adult learners of Spanish with L1 English (mixed)	Two short stories	56.4%	-	-
Qian (2005)	International students (high intermediate)	Short factual texts	45.0%	-	55.0%
Wesche & Paribakht (2010) (see also Paribakht 2005)	Iranian undergraduate students of English with L1 Persian, Canadian college students with L1 French (intermediate/advanced)	Six short general interest texts	21.0%	15.8%	63.2%

Table 3.8: Overview of lexical inferencing success in L2 reading studies adapted from van Zeeland (2014: 1007)

The studies summarized in Table 3.8 mostly used think-aloud procedures to capture lexical inferencing processes and found success rates ranging from approximately 20% to 60%. They indicate that greater proficiency and larger vocabulary sizes aid lexical inferencing, which is discussed in more detailed below. Several projects have also investigated the knowledge sources used to make inferences (Bengeleil & Paribakht 2004; Qian 2005; Wesche & Paribakht 2010). Findings in this area show that L2 readers tend to make more use of linguistic cues than of non-linguistic cues, such as world knowledge. In addition, intralingual (i.e. L2) cues are used far more often than interlingual cues involving the comparison of different languages. In terms of clue location, readers appear to mostly concentrate on the immediate sentence context of the target words.

What is of special interest here is the relation between successful lexical inferencing and gains in vocabulary knowledge, as some studies indicate that lexical inferencing does contribute to vocabulary learning. Using the VKS for pre- and post-tests, Bengeleil and Paribakht (2004) found a small, but significant increase in vocabulary knowledge, which was retained on a delayed post-test two weeks later. Similarly, Wesche and Paribakht (2010) found small, but significant gains on an immediate VKS post-test. Furthermore, both lexical inferencing success (r = .63) and acquisition of new word knowledge (r =.76) correlated significantly with participants' overall vocabulary size, but there was also a significant relationship between inferencing success and word learning (r = .73). Pulido (2007) found that 3.83 out of 8 target pseudowords were learned at the level of meaning recall and 6.13 out of 8 at meaning recognition. Passage sight vocabulary, i.e. the amount of words in the texts that were known by participants, had a significant effect on

the retention of inferred words. In addition, Pulido (2007) also found an impact of passage sight vocabulary size on the perceived ease of inferencing. Elgort (2017) specifically investigated the impact of correct or incorrect inferencing on contextual word learning using measures of both explicit and implicit vocabulary knowledge (see section 3.1.2). She found that participants were able to recall the meaning of 15% of the target words in a post-test, but the responses of those who had come up with incorrect inferences were about 7% less accurate. Interestingly, participants who had not managed to infer a meaning at all during the learning phase performed worse with 10% less accurate responses. Elgort (2017) also found a mediating effect of vocabulary size: for participants with larger vocabulary sizes the negative impact of incorrect inferences on word learning was much weaker. In sum, the findings of these studies show that lexical inferencing success depends at least partially on vocabulary size and is related to small, but significant gains in word knowledge. Furthermore, it indicates that incorrect inferences can have a negative effect on vocabulary acquisition, particularly for learners with smaller vocabularies.

Finally, a study that is worth having a closer look at is Haastrup (2008) because her sample includes learners similar to the participants in this study. Haastrup (2008) compared lexical inferencing success in both L1 and L2 reading in three groups of Danish learners of English: 30 students in 7th grade of comprehensive school, 30 learners in 10th grade of upper secondary school and 30 first-year students at university. With regard to lexical inferencing success in the L1, she found that the learners in 10th grade (50.1%) performed significantly better than the learners in 7th grade (28.9%) and the university students (58.8%) significantly outperformed the 10th grade learners. All groups scored significantly lower on the lexical inferencing task in L2 English (grade 7: 16.8%; grade 10: 37.3%; grade 13: 48.1%), but again the differences between the three groups were statistically significant. Haastrup's (2008) results show that lexical inferencing capability increases with age and educational level in both the L1 and the L2; therefore, these factors are likely to influence incidental learning in other learner populations as well.

This brief summary of research on lexical inferencing shows that L2 readers can guess words from context while reading and that it can contribute to incidental vocabulary learning as shown by small, but significant gains in word knowledge (Elgort 2017; Pulido 2007; Wesche & Paribakht 2010), although this is certainly not always the case. In addition, prior vocabulary knowledge was found to have an effect on lexical inferencing and word learning in these studies. Haastrup's (2008) study with Danish learners also highlights the effect of growing cognitive maturity, language proficiency, and vocabulary knowledge, as she found significant gains in lexical inferencing success with increasing age and educational level.

In sum, this section showed that contextual or incidental vocabulary learning from reading clearly is a reality, although it is a slow process as suggested, for instance, by Elgort and Warren (2014). In general, recognition knowledge is easier to acquire, as shown in a number of the studies reviewed, and participants are more confident about it than about recall knowledge

(Pellicer-Sánchez 2016). Frequency of occurrence was found to be a crucial influencing factor in studies using offline measures and eye-tracking methods (see section 3.1.3), although attention allocation as measured by total reading time in eye-tracking is a strong contender as well. Furthermore, findings in lexical inferencing research and the extraordinarily large results obtained by Pellicer-Sánchez (2016) suggest that language proficiency and prior word knowledge also play an important role.

3.3.2.2 *Vocabulary learning from listening*

In contrast to reading, few studies have investigated vocabulary gains from listening activities (Van Zeeland 2017), but two studies have aimed to compare the effectiveness of reading and listening for vocabulary learning. Brown, Waring and Donkaewbua (2008) conducted a study with three groups of Japanese university students who were presented with the same three stories containing 28 pseudowords in a reading, listening and reading-while-listening condition. Their findings show that the listening-only mode led to the smallest gains on a meaning recognition (29.3%) and recall measure (0.02% or 0.5 words) and that reading while listening led to the largest gains (47.5% for recognition and 15.7% for recall), while gains from reading only were somewhat lower (44.8% for recognition and 14.6% for recall). Statistical analyses showed that for both measures the difference between the listening-only mode and the other two modes was statistically significant. The authors thus conclude "that it is far more difficult to pick up words from listening-only than from either the reading-only or reading-while-listening modes" (Brown, Waring & Donkaewbua 2008: 148). Regarding retention of word knowledge, results of two delayed post-tests show that recall knowledge decayed much more than recognition knowledge and again highlights the role of frequency of occurrence: more than 50% of the target items that were met 15 to 20 times in the reading or reading-while-listening condition were recognized on the immediate post-test, but only approximately 30% in the listening mode.

Vidal (2011) compared vocabulary gains from reading and listening in an academic context: 248 first-year students in Spain were split into a listening group, a reading group and a control group. The two experimental groups encountered the same 36 target words either in three academic readings or three video-taped lectures, which closely corresponded to the reading texts. Using the VKS format, Vidal (2011) showed that reading led to significantly higher gains than listening across all knowledge dimensions measured by the VKS (i.e. knowledge of form, partial knowledge of meaning, precise knowledge of meaning, see Paribakht & Wesche 1997). However, the differences between reading and listening decreased with increasing English proficiency. Furthermore, it was found that frequency of occurrence was the influencing factor with the most predictive power for reading, while for listening it was predictability from word parts. Taken together, Brown, Waring and Donkaewbua's (2008) and Vidal's (2011) studies seem to indicate that reading is more effective for incidental vocabulary learning than listening, but that listening on its own can also lead to small vocabulary gains.

Early studies that concentrated solely on listening originated from research on modified oral input (e.g. Ellis 1995; Ellis & He 1999; Ellis & Heimbach 1997; Ellis, Tanaka & Yamazaki 1994;

Loschky 1994) and from research on story telling with children at pre-school level (e.g. Elley 1989; Feitelson et al. 1993). In addition, vocabulary learning through listening has also been investigated in instructed language contexts. Wode (1999), Donzelli (2007), and Horst (2010) compared possibilities for vocabulary learning from teacher talk or classroom discourse and coursebooks, Vidal (2003) explored vocabulary learning in academic lectures and Smidt and Hegelheimer (2004) used web-delivered lectures. While the studies on teacher talk are clearly related to school contexts, learners of English may encounter lecture-style input in extramural settings in certain YouTube videos or TED talks. However, they are more likely to use more informal sources of listening like podcasts (Putman & Kingsley 2009)[60] or songs (Schwarz 2013, see also Toffoli & Sockett 2014), which have been researched to a much lesser extent; thus, further studies using more stringent research designs would be desirable.

One study that investigated incidental vocabulary learning from listening to sources which learners could likely encounter outside school, is van Zeeland and Schmitt (2013a). Four short passages were taken from a TV talk show, TV interviews and an informal lecture and are linked by the common topic of crime. To measure incidental acquisition, 24 target items with different parts of speech were substituted with pseudowords. 30 postgraduate students listened to recordings of the adapted passages and were tested on their vocabulary gains on measures of form recognition, grammar recognition, operationalized as part of speech, and meaning recall. 20 participants took an immediate post-test and ten participants were tested two weeks after treatment. The results of the immediate post-test show that overall learning occurred for 7.05 of the 24 target words (29.2%). Form recognition knowledge was learned best (45.8%) followed by grammar recognition (33.7%), while meaning could only be recalled for 8.5% of the target items. The results of the delayed post-test show the same pattern but were significantly lower for form recognition (25%) and grammar recognition (24.6%), whereas there was no significant difference between immediate and delayed post-test for meaning recall (7.5%). Unlike the studies on reading (see section 3.3.2.1), this study did not find a strong effect for frequency of occurrence, but effects were found for the target items' concreteness and part of speech. Van Zeeland and Schmitt's (2013a) findings thus support the conclusions drawn above on the basis of the studies comparing reading and listening: it is unlikely that L2 learners will learn the meaning of new vocabulary from listening only, but it can be effective for establishing knowledge of form at recognition level.

Another study by van Zeeland (2014) is the first to explore lexical inferencing as another crucial aspect of vocabulary learning from listening.[61] In addition to comparing the relation between lexical inferencing success and several influencing factors among both L1 and L2 speakers, van Zeeland (2014) also examined the extent to which L2 learners notice unknown words in listening input. A total of 47 L1 users of English who were undergraduate students and 30 L2 users who were postgraduate students took part in the study. They listened to two listening passages with

---

[60] This study is one of few to investigate vocabulary uptake from podcasts and thus worth mentioning despite the focus on L1 science vocabulary.
[61] For an overview of all four studies in van Zeeland's inspiring PhD, see Van Zeeland (2017).

a more or less familiar topic (crime and malaria prevention), which each included 10 pseudowords. Participants' success at inferring the meaning of these pseudowords was measured in a one-on-one interview procedure. L2 users also completed a second task to measure their noticing ability: they listened to a recording containing pseudowords and had to press a button whenever they heard an unfamiliar word. Results concerning lexical inferencing show that L1 users were successful in 59.6% of the cases and L2 users in 35.6%. Both groups showed significantly higher inferencing rates when relying on local rather than on global cues and when they had higher background knowledge of the topic presented (see sections 3.1.3 and 3.3.2.1). Still, L1 users significantly outperformed L2 users, which is related to the fact that L2 users noticed the target items in the input far less often than L1 users: results of the second experiment show that L2 users only identified the pseudowords in the input in 44.2% of the cases. However, it seems that more advanced learners are better at noticing unfamiliar lexical items because there was a significant positive relationship between participants' receptive vocabulary size and their noticing ability. Interestingly, further analyses of the lexical inferencing data show that even when participants in the L2 group did notice the target pseudoword, they were significantly less likely to make a correct inference than those in the L1 group. Van Zeeland (2014) thus concludes that "[w]hile the results from this study suggest a lower success rate in listening than in reading, they also indicate that the general effect of the three variables [background knowledge, clue type and vocabulary size] is the same in both modes".

This section has presented evidence for the occurrence of incidental vocabulary learning from listening input. Although this area has received much less attention than incidental vocabulary learning from reading, the data accumulated to date suggest that vocabulary can be learned from listening, although it is less effective than reading, particularly for the acquisition of word meaning. However, much of the aural input learners of English receive in their extramural surroundings is supported by visual imagery; hence, the question whether audiovisual media are more beneficial for vocabulary learning than aural input only is addressed in the following section.

3.3.2.3 Vocabulary learning from viewing

Expanding on the previous section on listening, this section synthesizes research on audiovisual media, in which aural input is supported by visual imagery. This is the case in many popular EE activities such as watching video clips, films, series, or TV. Research on incidental language learning from viewing has mainly focused on vocabulary and has often been interlinked with the question of whether the use of L1 subtitles or L2 captions can support such learning (Danan 2004). The underlying premise that videos or subtitled videos are more beneficial to learning than audio alone stems from multimedia theories: Paivio's (1986, 2007) dual coding theory posits that information is processed more effectively when presented both verbally and visually and Mayer's (2014: 1) multimedia learning hypothesis similarly states that "people can learn more deeply from words and pictures than from words alone". While Mayer's theory does not focus

on second language learning, a study on L2 vocabulary acquisition by Bisson et al. (2013) has shown that even very brief exposures to multimodal input in an incidental learning phase lead to much higher gains in a subsequent explicit learning phase, thus "showing that informal exposure to multi-modal foreign language leads to foreign language vocabulary acquisition" (Bisson et al. 2013: 1). Research by Rodgers (2018) suggests, however, that not all video genres may support vocabulary learning equally well. He compared a fictional series and a documentary in terms of the visual support provided for concrete nouns in the audio track and found that in the documentary visual images occurred significantly more often in close proximity to the target nouns that in the narrative TV series.

As in the two examples mentioned so far, the vocabulary learning potential of audiovisual media has been investigated both in corpus-driven research and intervention studies. Corpus-based studies have helped to establish learning goals by investigating coverage levels for films and TV series (Webb & Rodgers 2009a, 2009b, see section 3.1.2), examined the potential for repeated encounters with low frequency words in movies (Webb 2010) and explored the possibility of acquiring formulaic sequences from watching TV (Lin 2014). The latter studies both found favourable conditions for vocabulary learning from extensive viewing (Webb 2015). Extensive viewing is recommended as Webb (2010) showed that few low frequency words occur a sufficient number of times (10 or more) to support incidental learning in a single movie. However, in his complete corpus consisting of 143 movies a large proportion of low frequency words occurs 10 times or more, thus pointing to the beneficial effects of regularly watching different movies. Lin (2014: 164) used the iTV corpus based on British television programmes to show that the occurrence of formulaic sequences in TV is "directly proportional to that in everyday speech" (as represented by the spoken component of the BNC), so that internet television can be regarded as an equally suitable source of learning for single and multiword items.

In addition to corpus-driven studies, intervention research has focused on learning gains from watching video input in different conditions. Early intervention studies (e.g. d'Ydewalle & Pavakanun 1995; d'Ydewalle & Van de Poel 1999; Koolstra & Beentjes 1999; Neuman & Koskinen 1992; Pavakanun & d'Ydewalle 1992) used short video clips with learners from various age groups and commonly compared a subtitling or captioning condition to a video-only condition. Several of these studies point to a superior performance of the group that was able to use subtitles or captions, but they also provide evidence that watching audiovisual input in any condition leads to incidental vocabulary learning. Newer studies have found additional support for these early findings (e.g. Lekkai 2014) and expanded on them in several ways.

Lin (2010) investigated the effects of reading and listening proficiency on incidental vocabulary learning from five short video clips based on CNN news. She used a multiple-choice pre-test and five post-tests (one after each viewing session) to compare the performance of 82 Chinese students in three proficiency groups. Her results indicate that two out of the three groups improved significantly over the course of the study, but the students with high listening abilities

and low reading abilities did not. Furthermore, the group of students who had good listening and reading skills at the outset of the study learned most lexical items, which is another example of the Matthew effect in language learning and also points to the necessity of including further independent variables to account for variation in results.

In an attempt to research vocabulary learning from watching online television in a more naturalistic way, Kusyk and Sockett (2012) compared the word knowledge of regular and non-regular viewers of English-language TV among 35 French students studying technology. They administered a questionnaire to participants and tested their knowledge of 42 highly common phrases identified in a corpus study by Sockett (2011) using a modified version of the VKS. They found a significant difference in knowledge of target phrases between students who reported regularly watching English-language TV and those who reported doing so only sometimes or rarely. While the approach to researching naturalistic learning is highly interesting, one problem of this study is that the phrases tested are highly frequent, conversational structures that the learners, who are at level B1 according to a self-evaluation measure, likely have also encountered in instructional settings.[62] In this light, the significant difference found between regular and non-regular viewers of English online television could be more of a practice effect and thus further studies are needed to explore naturalistic language uptake of novel words from audiovisual media.

A second attempt to investigate learning from video in a more ecologically valid way is a study by Arndt and Woore (2018), which also incorporates a direct comparison of vocabulary uptake from reading and viewing. A sample of 79 EFL learners from a variety of L1 backgrounds, either watched three video blogs (vlogs) on the topic of online friendship or read three corresponding blog posts.[63] Knowledge of six target pseudowords was measured at meaning recognition, meaning recall and form recall level in addition to recognition and recall tests of grammatical function. Mean total vocabulary gain across all knowledge aspects was 67% and there was no significant difference between the vlog and the blog group. Although the vlog group achieved higher scores for most aspects of word knowledge with small to medium effect sizes, the only statistically significant difference between the two groups was found for the form recall test, on which the blog group, who saw the orthographic form of the target words, outperformed the vlog group. The study thus provides "some tentative evidence" (Arndt & Woore 2018: 135) that multimodal videos may be more beneficial for learning aspects of word meaning and grammatical function than reading.

Criticizing that many previous studies used short video clips as input material, Rodgers (2013) is the first to use ten episodes of a TV series called *Chuck* in a longitudinal study spanning 13 weeks.[64] More than 200 Japanese first- and second-year university students watched one episode

---

[62] Examples include "What are you doing?", "I want you to", "What do you think?", "I don't want to", "What do you want" or "What the hell is" (Kusyk & Sockett 2012: appendix 2).

[63] If not explicitly stated otherwise, the studies reviewed in this section used English-language video material without captions or subtitles.

[64] A second study by Rodgers and Webb (2017) also used longer input materials in the form of ten episodes of a TV series; however, the focus of this article is on comprehension rather than vocabulary uptake. A second

per week and were then tested on 60 target items using two multiple choice tests. On average, the experimental group learned six new target items with a significant difference between pre- and post-test results. In addition, the experimental group also performed significantly better than a control group (N = 73). Rodgers (2013) also investigated the influence of frequency of occurrence and prior vocabulary knowledge on vocabulary uptake from watching the series. Surprisingly, he found no effect for prior vocabulary size, but a medium effect for frequency of occurrence.

A second study that used full-length input materials is Peters and Webb (2018), who researched incidental vocabulary uptake from one complete TV documentary. To avoid test effects, they conducted two experiments with L1 Dutch first- and second-year business students. A full-length BBC documentary on economy was used as input material and knowledge of 64 target items was tested at form recognition and meaning recall level in experiment 1 and at meaning recognition level in experiment 2. Participants in both experiments (experiment 1: N = 63, experiment 2: N = 62) were split in an experimental and a control group and completed a pre-test, post-test and delayed post-test one week after treatment. Due to the presence of a testing effect the results of the form recognition test were not analysed in detail, but experiment 1 showed that on average four target words were acquired at the level of meaning recall from watching the documentary. Results of the meaning recognition test in experiment 2 show that the video group knew 14% more target words on the post-test than on the pre-test. Moreover, statistical models found a significant effect for treatment in both experiments, indicating that the video groups outperformed the control groups concerning both meaning recall and recognition. The treatment variable accounted for 21% of the variance in scores in experiment 1 and 8% in experiment 2. In addition, the two experiments also investigated the potential impact of four influencing factors: frequency of occurrence, cognateness, prior vocabulary knowledge and relevance for understanding the input.[65] Significant effects were found for frequency of occurrence, cognateness and prior vocabulary knowledge in both experiments, with the last variable having the most impact and explaining 8% and 19% of variance respectively in experiments 1 and 2. The study thus clearly shows that vocabulary learning from audiovisual input results in vocabulary learning at both recall and recognition level and, unlike Rodgers (2013), the results indicate that participants' vocabulary size is an important influencing factor, which explained more variance than the treatment itself for the meaning recognition measure.

In addition to these studies on vocabulary learning from video input with or without subtitles, there is a second research strand that focuses on the use of L2 captions in combination with audiovisual input. A meta-analysis on the effects of captioning on L2 listening comprehension and vocabulary acquisition by Montero Perez, Van Den Noortgate and Desmet (2013) provides a useful overview of research in this area covering studies from 1989 to 2011. In their analysis concerning vocabulary-related research the authors found large effects for captioning on

---

publication focusing on vocabulary learning appears to be in preparation (see Rodgers 2018: 194), but cannot be included in this review at present.

[65] This variable is referred to as salience in section 3.1.3.

vocabulary learning as measured by immediate post-tests. These effects were present for recognition and recall knowledge. Interestingly, no impact was found for proficiency level with the effect size remaining stable across different groups. The authors conclude that their "[r]esults thus support the claim that captioning helps learners to improve comprehension and fosters vocabulary learning" (Montero Perez, Van Den Noortgate & Desmet 2013: 733).

Newer research generally also found positive effects for presenting audiovisual input in combination with L2 captions. Two studies by Montero Perez et al. (2014) and Montero Perez, Peters and Desmet (2017) found that groups who had access to different forms of captioning consistently outperformed video-only control groups, in particular with regard to the acquisition of word form (Montero Perez et al. 2014). However, a study by Jelani and Boers (2018) calls some of these results into question because of a potential bias effect of testing methodology: the authors argue that since earlier studies used written prompts in vocabulary tests, participants who saw captioned video, and thus the written word forms, may have been unfairly advantaged over participants who watched uncaptioned video. Hence, in their study with 16-year-old high school students in Malaysia half the prompts were delivered in written and half in auditory form in a form recognition and a meaning recall test. Their results showed a statistically significant difference between the caption and no-caption groups for meaning recall, but not form recognition. In addition, closer analysis of the meaning recall data showed that the significant difference was entirely due to the caption groups' superior performance on target items presented as written prompts. Jelani and Boers (2018: 169) thus conclude that further research "should take input-modality – test-modality congruency into account", although it is doubtful that the large effects found in previous studies depend on test modality alone.

Finally, a study by Peters, Heynen and Puimège (2016) is the first to directly compare the effects of L1 subtitles and L2 captions on initial vocabulary learning with secondary school learners and thus is of particular interest. In line with previous research, the authors hypothesized that captions would be more useful for the acquisition of word form, while subtitles would help learners to better understand the meaning of unknown lexical items due to the L1 translations. To test these hypotheses, they conducted two experiments with two different samples of learners in a general secondary and a vocational school. In the first experiment, 28 students in a general secondary school (17 to 18 years) watched a video clip on eating insects either with captions or subtitles. Knowledge of 39 target items was measured at form recognition and meaning recall level. On average, the students recognized 42.1% of word forms and recalled the meaning of 19.9% of target words. Type of subtitle was found to be a significant predictor in an interaction with prior word knowledge for form recognition, but not for meaning recall. The caption group thus performed better than the subtitle group regarding word form, as expected, but no effect of L1 subtitles was found on meaning recall. In the second experiment 18 lower-proficiency students attending a vocational school (17 to 20 years) watched a clip from the TV series *The Simpsons*. Knowledge of 18 target items was measured at form recognition, form recall and meaning recognition level. Overall, 27.4% of target items were known on the form recognition

measure, 16.2% for form recall and 24.2% for meaning recognition. Significant effects for type of subtitle were found for form recall, but not for the remaining two word knowledge aspects. The results of these two exploratory studies suggest that while L2 captions appear to have positive effects on the learning of word form, L1 subtitles do not have the hypothesized effect on the acquisition of word meaning. In addition, both prior vocabulary knowledge and frequency of occurrence were found to be mediating factors that need to be taken into account.

Taken together, these studies indicate that input material which integrates several modes of presentation such as images, speech and text is beneficial to vocabulary learning and probably more so than listening on its own, although to the best of my knowledge there are no recent studies directly comparing audio and video input. While there are several strands in this research area dealing with questions of captioning and subtitling, overall the studies show that watching audiovisual media such as online video clips, TV series or films leads to incidental vocabulary learning, even in video-only control groups. Audiovisual media are however not the only multimodal input that teenagers are likely to come across in their leisure time; hence, the next and final section of this literature review summarizes research on gaming.

3.3.2.4 Vocabulary learning from gaming and online environments

> As interactive, multimodal, immersive, and extremely popular environments, digital games have received increasing interest from educators in recent years for their potential to enhance the language learning experience, both inside and outside the classroom (Reinders 2017: 1).

In line with the statement above, gaming appears to be a fashionable research topic within SLA at the moment (see Reinders 2017 for an overview), and research on vocabulary learning seems to be particularly prominent within this area (Jabbari & Eslami 2019). This is also indicated by the fact that two meta-analytic studies recently synthesized research on game-based vocabulary instruction and learning (Chen, Tseng & Hsiao 2018; Chiu 2013). Both found positive effects of digital games on vocabulary learning; however, many of the studies included in these reviews concern the integration of games into teaching, whereas this section is not concerned with gaming in instructed contexts. Moreover, studies designing, evaluating and or using educational vocabulary games are also excluded because teenagers are unlikely to spend much of their free time with pedagogical games. To better reflect this project's focus on EE activities, the studies included in this section focus on commercial off-the-shelf games.

Gaming is potentially beneficial to language learning because of the integration of multiple modes, similar to audiovisual media (see section 3.3.2.3). However, even more importantly, it is one of the few EE activities in which learners actively interact with and through L2 English. Understanding the language is crucial to fulfilling quests or solving puzzles and therefore there is an intrinsic motivation to continuously improve one's language proficiency.[66] Interactivity is indeed the defining characteristic of digital games: first, learners have to interact with the

---

[66] This fact is most dramatically expressed in the title of an article by Reinders and Wattana (2010): "Learn English or die: The effects of digital games on interaction and willingness to communicate in a foreign language", which however, as the title says, focuses on oral interaction rather than vocabulary learning.

gaming environment itself in order to be successful. Second, in MMORPGs (e.g. *World of Warcraft*) or in games providing the option to play in online teams (e.g. *Counterstrike, League of Legends*) gamers frequently also have to interact with other players (see also Sundqvist 2013). Since online gaming communities are frequently international, English often functions as a lingua franca between gamers from different L1 backgrounds (Bytheway 2015; Sylvén & Sundqvist 2012a). Digital games, and in particular multiplayer online games, thus appear to provide rich ground for vocabulary development in general and for English in particular.

Research has addressed the question of whether and how vocabulary is learned from gaming from several angles. Qualitative studies based on interviews and/or observations have focused on learners' perceptions of learning from digital games (Turgut & İrgin 2009), the use of VLS (Bytheway 2015), individual learning histories (Sundqvist 2015), or dialogical learning (Zheng, Bischoff & Gilliland 2015). These qualitative studies show that learners are highly motivated to understand the language used in the game (Bytheway 2015; Sundqvist 2015; Turgut & İrgin 2009), use a variety of vocabulary learning strategies (Bytheway 2015; Turgut & İrgin 2009) and find lexical repetition in in-game texts helpful (Turgut & İrgin 2009). In addition, this research also suggests that the meaning of new lexical items can be acquired through interaction with other players (Zheng, Bischoff & Gilliland 2015) and that gamers give and receive language-related explanations and feedback (Bytheway 2015; Sundqvist 2015).

In contrast, more quantitatively oriented studies have typically tried to measure vocabulary uptake from gaming in general (e.g. Hannibal Jensen 2017; Sundqvist & Wikström 2015; Sylvén & Sundqvist 2012a) or from playing specific computer games (e.g. Chen & Yang 2013; Rankin, Gold & Gooch 2006). In a small-scale study focusing on the game *Ever Quest 2,* Rankin, Gold and Gooch (2006) found that after 4 weeks of gameplay participants knew the meaning of at least 35% of the words which occurred once in the interaction with non-playing characters, but 55% or more of the words that were used more than five times in these chat conversations, which again points to an influence of frequency of occurrence. Vocabulary uptake in Chen and Yang's (2013) study of the adventure game *BONE* with 22 Taiwanese students of intermediate English proficiency was smaller: after 1.5 hours of gameplay participants had acquired 2 out of 20 target words at meaning recall level.

Since interactivity is a defining characteristic of digital games, studies that explore the influence of this feature on vocabulary uptake are of particular interest. The first to look into this issue were DeHaan, Reed and Kuwada (2010), who compared vocabulary gains among two groups of Japanese undergraduate students of computer science: half of the 80 participants played the game *Parappa the Rapper 2*, in which players have to complete rap lines, for 20 minutes, whereas the other half only watched the players. Both groups were explicitly told to try and remember the lyrics used in the game and were then tested on their form recall knowledge of 41 target words using a gap-fill measure of the lyrics presented in the game immediately after treatment and two weeks later. Results of the immediate post-test showed that surprisingly players recalled significantly fewer words than watchers and the same result was also found on the delayed post-

test. However, a more recent study by Mohsen (2016) seems to contradict DeHaan, Reed and Kuwada's (2010) initial findings. 43 Arabic-speaking second-year students of English were divided into two groups, with one group playing a simulation game of a surgery, while the other group only watched a video of the simulation. Knowledge of 14 target items was measured at recognition level in an image identification task. Mohsen (2016) found that both groups had gained some knowledge as they knew more words on the immediate post-test than on the pre-test, but the players correctly identified significantly more target items than the watchers. An important difference between DeHaan, Reed and Kuwada (2010) and Mohsen's (2016) study is that unlike in the former the watchers did not actually watch the players, but saw a video instead. Mohsen argues that because of this set-up watchers in his study could not learn from the errors of other participants and could not avoid them accordingly, as they may have done in the earlier study. More importantly, the two studies also measured vocabulary acquisition at two different levels of mastery, recall and recognition, which could have had a decisive impact on the findings. Overall, the evidence as to whether interactivity as a key characteristic of digital games is beneficial to vocabulary learning is inconclusive at present, although the correlational studies presented below seem to suggest a relation between gaming and vocabulary knowledge.

The remainder of this section focuses on research carried out in Europe, which allows insights into gaming practices of children and teenagers in different age groups. All four studies have been carried out in Sweden and Denmark, where English plays a large role from an early age onwards due to subtitling practices (see section 1.1). Beginning with the oldest group of learners, Sundqvist and Wikström (2015) report on the gaming practices of Swedish teenagers in grade 9 of secondary school (15 to 16 years).[67] Based on time spent with digital games, they identified 35 non-gamers, 26 moderate gamers, who played an average of two hours a week, and 19 frequent gamers, who all played more than five hours a week but reported an average gaming time of almost 14 hours. The groups show a large gender difference with 95% of the frequent gamers being male and 89% of the non-gamers being female. Using statistical tests, Sundqvist and Wikström (2015) show that the frequent gamers outperform the other two groups in terms of receptive and productive vocabulary knowledge as measured by adapted versions of the VLT and PVLT. Advantages for the frequent gamers in terms of lexical knowledge were also found in an analysis of essays produced by the students. Sundqvist and Wikström (2015) thus conclude that their "findings indicate a positive relation between gameplay and L2 English - at least for boys".

Also in Sweden, Sylvén and Sundqvist (2012a) collected data on gaming among 86 students in grade 5 (11 to 12 years) and examined links to vocabulary knowledge as well as reading and listening skills. EE data were gathered using a questionnaire and language diary, while vocabulary knowledge was measured using a self-designed test covering meaning recognition at the 1K and 2K level and productive recall knowledge at the 2K level. Again, boys were found

---

[67] The data are taken from Sundqvist's (2009a) study, but the authors conduct a post-hoc comparison of three groups of gamers and their performance on the vocabulary measures used in the original study.

to spend significantly more time gaming than girls (4.4 vs 1.1 hours a week on average), but the questionnaire data suggests there is also a difference with regard to the types of games played: whereas boys frequently mentioned MMORPGs and first-person shooter games such as *World of Warcraft*, *Call of Duty*, or *CounterStrike*, the girls appear to prefer single-player simulation games such as *The Sims*, *Restaurant City*, or *Zoo Tycoon*. To analyse a potential positive relationship with vocabulary knowledge, participants were again divided into three groups of frequent gamers (*N* = 17), moderate gamers (*N* = 40) and non-gamers (*N* = 29). Similar to Sundqvist and Wikström (2015), over three quarters of the frequent gamers were boys. Furthermore, significant differences were found between the three groups with regard to vocabulary knowledge with frequent gamers performing better than moderate gamers, who, in turn, had higher scores than non-gamers. In a further study on young Swedish learners Sundqvist and Sylvén (2014) collected EE data from 76 pupils in grade 4 (10 to 11 years) and found results very similar to Sylvén and Sundqvist (2012a) concerning gaming behaviour, although this study did not focus on vocabulary knowledge. Already at this young age, frequent gamers, who were again mostly male, spend a mean time of 6.6 hours a week playing games in English, showing that already young gamers are exposed to a considerable amount of language input via games, at least in the Swedish context.

A study by Hannibal Jensen (2017) carried out in Denmark focused on even younger language learners. Due to a previous change in the curriculum the two groups of learners aged 8 and 10 years both only had received one year of English instruction: 49 were early starters, who had started English in grade 1, and 58 were late starters, who had begun English in grade 3. To investigate the relation between gaming and vocabulary size, Hannibal Jensen (2017) used the PPVT-4 and a language diary, in which students reported on their EE activities with the help of their parents. Similar to other EE studies (see section 3.3.1), she found that even young learners spend a considerable amount of time with English outside school with a mean EE time of over six hours per week. Participants spent more time with gaming than with watching audiovisual media or listening to music, but again, there was a significant difference between boys and girls. In addition, students hardly ever used Danish for playing digital games; thus, even for 8- to 10-year-olds English is the main gaming language. Vocabulary knowledge was tested using a pre- and post-test, which were administered one year apart. The mean total PPVT post-test score was 56.8 (*SD* = 25.7) with late starters performing significantly better than early starters. Findings concerning the correlation between games and vocabulary are presented according to type of game played and show that vocabulary outcomes and games with spoken and written English input are significantly related (r_τ = .218). Hence, Hannibal Jensen's (2017) study provides further evidence that gaming can contribute to vocabulary learning already at a very young age.

Most recently, Sundqvist (2019) investigated the relationship between time spent with digital games, type of games played and vocabulary knowledge with a quantitative sample of 1,069 Swedish learners in grade 9 and a qualitative sample of 16 learners in the same age group. The same shortened and adapted versions of the VLT and the PVLT as in Sundqvist (2009a) were

used to test vocabulary knowledge in the quantitative sample and a questionnaire was employed to gather data on gaming. In the qualitative sample, an essay that forms part of a national assessment was collected and semi-structured interviews were conducted with pairs or small groups of students. Sundqvist (2019) found that frequent gamers scored highest on both vocabulary measures and that time spent playing commercial off-the-shelf games correlated significantly with both the VLT (r_s = .31) and the PVLT (r_s =.28). In addition, a relationship between time spent gaming and type of game was identified: players who preferred single-player games tended to spend less time gaming, while those who played multi-player games or MMORPGs spent significantly more time with digital games. With regard to vocabulary size, the latter group of gamers had significantly higher mean scores than low-frequency gamers with a preference for single-player games, or non-gamers. Analysis of the essays collected from the smaller qualitative sample generally support the conclusions based on the quantitative data according to Sundqvist (2019).

The studies summarized in this section provide evidence that playing digital games has beneficial effects on L2 vocabulary learning. Correlational studies from Sweden and Denmark show that gaming is a popular EE activity among 8- to 16-year-olds and that a considerable amount of time is spent playing games in English. However, this finding is mostly true for boys, as the studies identified significant differences between girls and boys in relation to time spent gaming (Sylvén & Sundqvist 2012a; Hannibal Jensen 2017) and qualitative differences in relation to the types of games played (Sylvén & Sundqvist 2012a). In addition, intervention studies suggest that new vocabulary is learned incidentally from playing digital games (Rankin, Gold & Gooch 2006; Chen & Yang 2013). Findings concerning the key feature of interactivity are inconclusive at present, but recent research by Mohsen (2016) points towards beneficial effects. Finally, qualitative data show that gamers find visual imagery helpful for understanding new vocabulary items, receive support from other players, and benefit from repetition in the game or repeated playing.

In addition to gaming, other forms of online interaction, and in particular online writing in the form of blogs, forums, fan fiction or even social media, could potentially also contribute to language development in general and vocabulary learning more specifically. Surprisingly, I have not been able to find research on vocabulary learning in relation to online contexts like these, which may be related to the overall scarcity of studies on vocabulary learning through output (Nation & Webb 2011). While there are studies on the development of vocabulary use in writing (e.g. Crossley et al. 2011; Dóczi & Kormos 2016), little research has investigated vocabulary learning through engagement in writing tasks. Barcroft (2004), Webb (2005) and Webb and Piasecki (2018) look at the effects of writing in intentional learning and Nation and Webb (2011) mention an unpublished dissertation by Coxhead (2008, as cited in Nation & Webb 2011: 127), but to the best of my knowledge there are no studies with an explicit focus on vocabulary learning in fan fiction communities and other online contexts involving written output and/or social interaction.

This research gap seems surprising since applied linguistic studies exploring online interest groups and fan fiction communities (e.g. Lam 2000; Thorne 2009), and in particular studies by Black (2005, 2008), suggest that "online fan fiction communities promote informal, participatory types of learning that are beneficial for adolescents' L2 literacy development, as well as for their sense of self-efficacy in, and level of affiliation with, English" (Thorne, Black & Sykes 2009: 805). Further research shows that imitation and feedback are vital elements in fan fiction communities (Olin-Scheller & Wikström 2010) and that mentoring takes place at several levels (Campbell et al. 2016; Evans et al. 2017). Isbell (2018: 91) found that members in an online community specializing in Korean mostly focused on "asking questions or sharing knowledge about grammar, vocabulary, and pronunciation". These studies indicate that online communities sharing a special interest offer unique opportunities for informal language learning, particularly if they are focused on literacy, as is the case with fan fiction, or on learning languages. Proposals have been made to include such community practices in language teaching (e.g. Sauro 2017) and the potential of such online spaces, in addition to gaming or viewing, has recently been highlighted specifically in relation to contextualized vocabulary leaning (Godwin-Jones 2018). So far, however, this line of investigation seems not to have been taken up in empirical vocabulary research.

Taken together, the studies reviewed in this section point to several important findings: first, research summarized in section 3.3.1 indicates that learners of all age groups studied so far engage with extramural English. Second, English input outside school is mostly consumed via popular media such as music, TV programmes or films, and games, and often involves a screen. Third, research suggests a positive relationship between such informal engagement with English and vocabulary development and shows that many young learners acquire knowledge of English in this way before the start of formal instruction. Fourth, although many studies were conducted in European countries where English television with subtitles constitutes a major source of input, evidence on the effects of TV is mixed, which indicates that subtitled TV programmes are not the decisive source for language and vocabulary learning from EE. Fifth, some studies even indicate that the amount of engagement with EE may show a stronger relationship with learning outcomes than length and quantity of formal instruction at school.

Research on vocabulary learning from different activities presented in section 3.3.2 suggests that reading, listening, viewing and gaming can contribute to incidental vocabulary acquisition. Overall, recognition knowledge is easier to acquire from all activities than recall knowledge. Gains are typically larger for reading than for listening, while direct comparisons of reading and watching video are rare. Both audiovisual media and digital games support vocabulary through multimodal input, which can be further enhanced through the use of L1 subtitles or L2 captions, especially the latter have been found to be beneficial for the acquisition of form. For games, interactivity is a second defining feature because gaming is the only extramural activity in which learners have to interact with language in the game and frequently also with other players. With

regard to such in-game interactions, English has been found to be the predominant lingua franca among players; thus, communication between players also offers learning opportunities.

Concerning influencing factors, frequency of occurrence appears to play a greater role in reading than in listening but was also found to affect learning from watching audiovisual media. Prior vocabulary knowledge also emerges as a factor affecting vocabulary learning in many studies, particularly with regard to lexical inferencing, which can be seen as an initial step towards vocabulary acquisition. In relation to digital games, the evidence concerning the role of interactivity is inconclusive at present, but the most recent research points to beneficial effects.

3.4   Summary

This chapter provided information on the foundations of vocabulary research, vocabulary learning and vocabulary measurement and synthesized studies investigating the relationship between vocabulary knowledge and extramural English as well as incidental learning from related activities such as reading, listening, watching audiovisual media and gaming.

First, fundamental issues such as conceptualizations of a 'word' and of 'knowing' a word were discussed because the diversity of approaches found in the field has important implications for vocabulary testing and the interpretation of the results of empirical research. Next, our current state of knowledge regarding L2 vocabulary development was reviewed, highlighting the incremental nature of the learning process, the role of frequency, and the complementarity of incidental and intentional learning, which are both necessary to reach the aims of vocabulary learning identified in coverage research. Learning outcomes in relation to vocabulary size found in recent empirical studies with European learners were summarized to contextualize the findings of the present study and factors influencing lexical learning from input were briefly set out to provide a background to the studies on incidental vocabulary acquisition.

Referring back to the lack of agreed upon definitions in the field, the discussion of test constructs and processes of operationalization in vocabulary size testing showed that measuring breadth of vocabulary knowledge in a valid, reliable and useful way is everything but an easy task. A range of decisions regarding the type of word knowledge to be tested, the unit of counting, the sampling of target items, the test format, the scoring procedures, and the presentation and interpretation of results need to be taken, ideally guided by an explicitly stated test purpose. In addition, evidence for the validity, reliability and usefulness of any vocabulary measure should be provided by test developers, which is currently often not the case. Guidance should also be provided for test users who employ vocabulary size tests for pedagogical and research purposes to avoid misuse. Adaptations of existing measures may affect research outcomes in undesirable ways, but due to their apparent simplicity, which stands in stark contrast to the complex nature of the decisions behind vocabulary size tests, this does not always appear to be clear to researchers using these tests in empirical studies.

In a third step, this chapter zoomed in more closely on the object of study in the present project and presented previous research on the relation between extramural English and vocabulary

knowledge and learning in diverse contexts. Overall, these studies point to a positive relationship between engagement with English in informal contexts and vocabulary knowledge, in particular vocabulary size, for diverse groups of learners at primary, secondary and tertiary level. Learners of all ages frequently engage with extramural English through popular media and thus have previously unprecedented opportunities for (incidental) vocabulary learning in their leisure time. Research focusing on vocabulary learning from specific activities like reading, listening, viewing and gaming provides more detailed information on the acquisition of different word knowledge aspects and the learning gains that can be expected, which are typically rather small for incidental learning. Reading generally emerges as more beneficial to lexical learning than listening, but recent studies show that vocabulary can also be learned from watching audiovisual media and playing digital games.

4 The research context: English in Austria

This chapter provides an overview of the research context of the present study, moving from the wider setting in Austria to the specific phenomenon of extramural English within it. After a brief introduction to the linguistic situation in Austria (section 4.1), information on the school system in general and English teaching in particular is given (section 4.2.1). Next, the present use and positioning of English in the public sphere is discussed (section 4.2.2) before considering relevant aspects of adolescents' lives in Austria (section 4.3). Finally, a characterization of the specific context of this study is presented after a summary of previous research on extramural English in Austria (section 4.4)

4.1 The linguistic situation in Austria

The linguistic situation in Austria is far more complex than is apparent at the level of official language policy. As the most important policy document, the Austrian Constitution (Bundes-Verfassungsgesetz) positions German as "Staatssprache der Republik", the state language of the republic (art. 8, section 1). In addition, the constitution grants legal rights to autochthonous ethnic minorities (art. 8, section 2) and to the Austrian sign language (section 3), which was recognized as an independent language in 2005 (Österreichischer Gehörlosenbund n.d.). The six autochthonous minorities include the following ethnic groups: Burgenland-Croats, Czechs, Hungarians, Roma, Slovaks and Slovenians (Bundeskanzleramt Österreich 2018). In addition to other protective measures, the languages spoken by these minorities can be used for official business with public authorities and are subject to special regulations regarding education and cultural promotion (Parlamentsdirektion n.d.).[68]

While these are the only legally recognized minority languages, in practice, there is much greater multilingualism in Austria than suggested by the national legislation. A report by the Austrian Integration Fund (Vasilyev 2012) on migrant languages in Austria includes 17 additional languages based on census data from 2011 and asylum statistics for the years 2009 to 2011. In the order of the approximate size of their speaker communities these are: (German) German, Serbian, Bosnian, Croatian, Turkish, Romanian, Polish, Russian, Chechen, Dari (Eastern Farsi), Pashto, Persian (Western Farsi), Arabian, Kurdish, Somali, Georgian and Albanian. Although exact speaker numbers are hard to determine (Vasilyev 2012), this list of languages is a testament to the growing linguistic diversity of the Austrian population. More recent asylum statistics for the year 2017 show that since the publication of Vasilyev's (2012) report the situation has further evolved: recent migratory moves have led to even greater diversification since the nations from which most asylum seekers filed applications in 2017 were Syria, Afghanistan, Pakistan, Iraq, Nigeria, the Russian Federation, Iran, Somalia, and Ukraine

---

[68] In terms of schooling, there are provisions for Slovenian schools in Carinthia and (Burgenland-)Croatian as well as Hungarian schools in Burgenland, granting more legal rights to these three minorities than the other autochthonous groups. In addition, there are (private) initiatives for Czech and Slovak schooling in Vienna, while there appears to be only one school in Burgenland offering lessons in Romani (Fischer & Doleschal 2013).

(Bundesministerium für Inneres 2017: 6). However, despite the ever-growing multilingualism of the population, the few existent official language policies solely concentrate on the long-standing autochthonous minorities described above, extending no explicit legal rights to widely-used migrant languages.[69]

Indeed, the lack of official regulations for migrant languages is just one example of a general absence of explicit language policies in Austria, which are rare outside education (see section 4.2.1). Several authors (De Cillia 2003, Dorostkar 2014: 136–155) have criticized and lamented the fact that, with the exception of provisions for autochthonous minority languages and regulations concerning language teaching at educational institutions, Austrian language policy is characterized by a non-interventionist approach, which only reacts in case of conflicts.[70] As a consequence, verbal, the Austrian section of the International Association of Applied Linguistics (AILA) passed the following verdict in the *Klagenfurter Erklärung 2011* [Klagenfurt declaration 2011]: "In Österreich fehlt derzeit eine koordinierte und längerfristig geplante Sprachenpolitik, die den gesamteuropäischen Zielsetzungen entspricht"[71] (De Cillia & Vetter 2013: 342). This lack of explicit policies does, however, not only concern minority or foreign languages, but also the standard itself: Soukup and Moosmüller (2011: 40) note that with the exception of the *Österreichisches Wörterbuch* [Dictionary of Austrian German] "official language policy-making that might publicly establish and define an Austrian standard German is virtually non-existent" (see also De Cillia 1997, 2003).

In sum, this brief overview of the linguistic situation in Austria shows that despite the small size of the country (8.8 million inhabitants in 2018, Statistik Austria 2020) there is great and growing linguistic diversity, mainly due to immigration. This increasing diversification is however not met by concomitant changes in official language policy and explicit linguistic regulations remain scarce. One of the few domains which is subject to top-down language policies is education; these will be described in the next section when considering the role of English in Austria.

---

[69] In fact, the only policy taking migrant languages into account is found in the educational context: so-called "mother tongue teaching" (*muttersprachlicher Unterricht*) is offered to all students who mainly or partly use a language other than German at home. These are optional subjects (*Freifach* or *unverbindliche Übung*) mainly provided by primary and lower secondary schools, in which participation is voluntary (see De Cillia & Haller 2013: 145, Fleck 2013). In the school year 2015/16 L1 teaching for 26 languages was offered (Albanian, Arabic, Bosnian/Croatian/Serbian, Bulgarian, Chechen, Chinese, Czech, Dari, French, Greek, Hungarian, Italian, Kurdish (Kurmanji), Nepali, Pashto, Persian, Polish, Portuguese, Romani, Romanian, Russian, Slovak, Slovenian, Somali, Spanish and Turkish), with most students attending classes for Turkish or Bosnian/Croatian/Serbian (Bundesministerium für Bildung 2016/17: 11). It does however not mean that all of the languages listed were taught nationwide and that all students participate in such classes. In reality, statistics show that in the same school year only 26.6% of all eligible primary school pupils took part in mother tongue lessons and between 3.1% and 11.2% at the different types of lower secondary schools, resulting in a proportion of 16.1% across all school types (Bundesministerium für Bildung 2016/17: 9). This is not necessarily a result of a lack of interest on the part of the students because these optional subjects can only be taught if a certain number of participants registers for them (Bundesministerium für Bildung und Frauen 2015/16): usually, a minimum of 8 students is required for the autochthonous minority languages (see footnote 68) and 12 for all other languages.

[70] The most famous language political conflict in recent history is the so-called *Ortstafelstreit* about German-Slovenian bilingual signs for place names in Carinthia, which lasted for almost 40 years (see Dorostkar 2014: 139).

[71] "In Austria there is currently no coordinated and long-term language policy that corresponds to pan-European objectives" (author's translation).

4.2   The role of English in Austria

Both from a sociolinguistic point of view (e.g. Kachru's (1985) model of World Englishes) and in the Austrian self-conception, English in Austria has traditionally been considered a foreign language. Indeed, English has no official status in Austria. Still, the label of 'foreign' language is questionable nowadays as its position is very different from other foreign languages, as we shall see in the following. In the latest Special Eurobarometer on languages (European Commission 2012b) 73% of all Austrian participants stated that they speak English well enough to hold a conversation, six times more than for French as the language in second place. In addition, more than three quarters named English as a language they considered useful for their personal development (European Commission 2012a). With this indication of widespread knowledge and perceived usefulness in mind, this section explores the role of English in Austria, beginning with the educational context before moving to the wider public sphere.

4.2.1 English in education

In the following, a brief description of the Austrian education system is given before turning to the teaching of English and other foreign languages within it. Austrian schools at all levels except for pre-primary education are organized at federal level, meaning that they are subject to the School Organization Act (*Schulorganisationsgesetz*) and that all curricula are provided by the Austrian state (De Cillia & Krumm 2010: 154). As is graphically depicted in Figure 4.1, children enter the formal school system at the age of six, which marks the beginning of compulsory schooling, and attend primary school for four years (grades 1-4). At the beginning of lower secondary school the ten-year-olds are separated into two streams: they either attend the lower level of an academic secondary school (*Allgemeinbildende Höhere Schule, AHS*) or a new secondary school (*Neue Mittelschule, NMS*).[72] Pupils attending the latter school type receive further education for four years and have then almost completed their compulsory schooling, which ends at the end of grade 9. If students choose to leave the school system after grade 9 to pursue an apprenticeship, their final year of compulsory education is usually completed at a pre-vocational school (*Polytechnische Schule*). However, pupils from a new secondary school may also transfer to an upper secondary school, both vocational and academic, although the latter is relatively uncommon (see Statistik Austria 2018: 50).

---

[72] In practice, the general secondary schools (*Hauptschule*) shown in Figure 4.1 do not exist anymore since the school year 2015/16 following a decision by the government to introduce new secondary schools in 2008/09 (see Statistik Austria 2018: 26, European Commission/EACEA/Eurydice 2018). The early age at which students have to decide which educational track they want to follow is frequently criticized; for instance, a 2016 report by the OECD states that "[s]tudent's socio-economic background has a key impact on their achievement and educational trajectory through Austria's stratified school system that is characterised by early tracking and selection" (Nusche et al. 2016: 13), thus arguing that children from families with lower SES as well as those with an immigrant background are disadvantaged in this dual system.
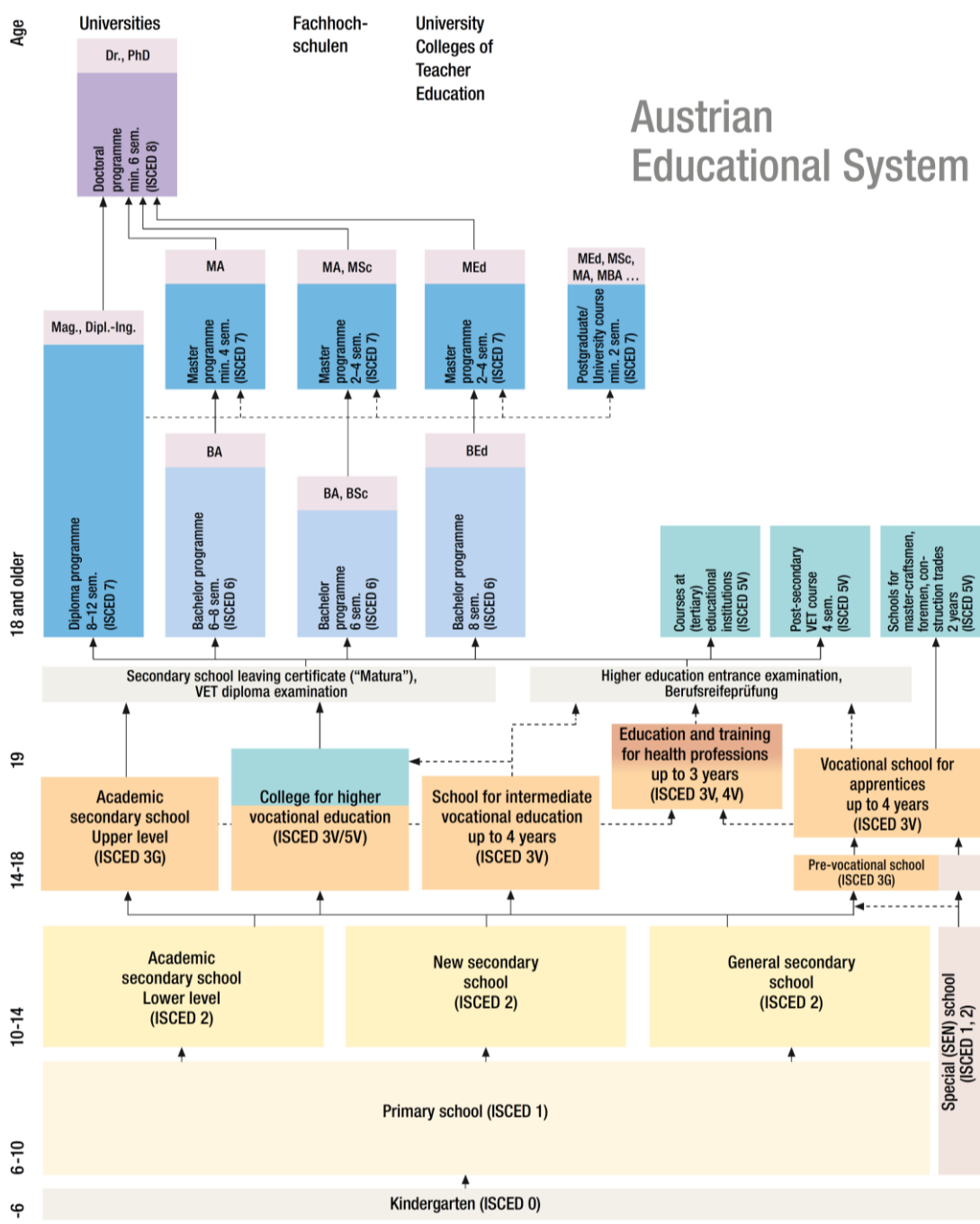
Figure 4.1: Schematic representation of the structure of the Austrian educational system provided by the Federal Ministry of Education (2017)

In contrast, pupils entering academic secondary schools (AHS), the so-called *Gymnasium*, at the age of ten can stay at the same institution for the eight years of their secondary education, transitioning to upper secondary school after grade 8 at the age of 14 and taking their A-levels after a further four years in grade 12. They can, however, also decide to leave and pursue their upper secondary education at a vocational school or college. Hence, both students who have completed a new secondary school (NMS) and those who have attended the lower level of an academic secondary school (AHS) can attend vocational schools from grade 9 onwards. There are many different types of vocational schools in Austria: one distinction is between schools for intermediate vocational education (*Berufsbildenden Mittleren Schulen, BMS*), which last three to four years, and colleges for higher vocational education (*Berufsbildenden Höhere Schulen, BHS*), which last five years and end with A-level examinations in grade 13. In addition, vocational schools offer many different types of specializations ranging from technology and engineering over trade and commerce to tourism.

At the end of their secondary education students at both academic and vocational upper secondary schools (AHS and BHS) take the Austrian school-leaving examinations, the so-called *Matura.* These present university entrance qualifications and have recently undergone major changes: a new standardized and competence-oriented exam (*Standardisierte Reife- und Diplomprüfung*) has been introduced in all Austrian upper secondary schools starting with the school year 2014/15 in academic schools (AHS) and one year later (2015/16) in all vocational colleges (BHS) (Bundesministerium für Bildung, Wissenschaft und Forschung n.d.-b). The written exams are provided by the Ministry of Education, Sciene and Research (*Bundesministerium für Bildung, Wissenschaft und Forschung BMBWF*) for the language of schooling (German, Croatian, Slovenian and Hungarian), modern foreign languages (English, French, Italian, Spanish), classical languages (Latin, Greek), mathematics and applied mathematics (Bundesministerium für Bildung, Wissenschaft und Forschung n.d.-b). The oral examinations follow a pre-determined structure, despite the tasks still being set by the class teachers.[73]

Following this brief overview,[74] it is worth providing a few more details on the academic stream of the Austrian school system as all participants in this study attend the upper secondary level of academic secondary schools (AHS). According to the curriculum (Bundesministerium für Unterricht und kulturelle Angelegenheiten 2000) students in these schools should be provided with a comprehensive and profound general education and be prepared for university entrance qualifications. In the school year 2016/17 35.5% of all pupils transitioning from primary to secondary level entered a lower academic secondary school (113,780) and a total of 91,906

---

[73] In addition to written and oral exams, the third pillar of the new standardized school-leaving exam is a "pre-academic paper" (*Vorwissenschaftliche Arbeit*) at academic secondary schools (AHS) and a "diploma paper" (*Diplomarbeit*) at vocational colleges (BHS). These papers are written during the last two school years of secondary education and are meant to prepare students taking their A-levels for further education.

[74] For further information on the Austrian educational system see Statistik Austria (2018), Austrian Agency for International Cooperation in Education and Research (OeAD) (2014) or European Commission/EACEA/Eurydice (2018).

students attended an upper academic secondary school, 57% of which were female (Statistik Austria 2018: 25–26). The larger proportion of girls at upper secondary level can be explained by the fact that more boys than girls change to colleges for higher vocational education after having completed the lower level of an AHS (Statistik Austria 2018: 50).

Turning to foreign language teaching within the Austrian school system, I will provide a brief account across school types before focusing on academic secondary schools. In primary education, compulsory foreign language teaching was introduced in grades 3 and 4 in 1983 and since 2003 it has become obligatory from grade 1 onwards (De Cillia & Haller 2013: 150). While in theory it is possible to teach all recognized minority languages (Croatian, Czech, Hungarian, Slovak, Slovenian) and neighbouring languages (e.g. Italian) in addition to English and French, in practice almost 99% of primary school pupils study English (De Cillia & Haller 2013: 159).[75] Virtually all Austrian children thus receive some form of English teaching from the age of six onwards, and some even learn English at pre-primary level (Buttaroni 2013). One could therefore assume that by the time they reach secondary education the pupils have acquired basic skills in English, which can then be built upon. In reality, there is great variation in the quality of English teaching at elementary level and thus in the level of achievement. Buchholz (2007) found large discrepancies between the educational policies for the provision of foreign language teaching in Austrian primary schools and the realities in classrooms, which is mainly due to the fact that primary school teachers, who are general education teachers and commonly instruct their classes in all areas of the curriculum, were assigned an additional task without the necessary training and little to no support (see Buchholz 2007: 321–323).

At lower secondary level the picture is similar to primary level in that the curricula offer a wide range of possibilities for foreign language teaching: Bosnian/Croatian/Serbian, Czech, English, French, Hungarian, Italian, Polish, Russian, Slovak, Slovenian, and Spanish can be taught in both streams of lower secondary education (AHS and NMS), while new middle schools can also offer Turkish (see Bundesministerium für Unterricht und kulturelle Angelegenheiten 2000, Bundesministerium für Unterricht, Kunst und Kultur 2012). In practice, the most common foreign language taught is again English with 98.8% studying it at this level. Furthermore, most students (89.8%) only study one foreign language at lower secondary school (De Cillia & Haller 2013: 158–159), meaning that for the majority of 10- to 14-year-olds English is the only language they study in addition to German.

More linguistic diversity is found at upper secondary level: about a third of the students (32.4%) study two and 4.2% study three foreign languages (De Cillia & Haller 2013: 160). However, again not much of the theoretically possible plurality of languages is translated into practice: English is the dominating foreign language followed by French, Italian and Spanish. All other languages

[75] The data reported in de Cillia and Haller (2013) and other publications (e.g. Nagel et al. 2012, De Cillia & Krumm 2010) are based on the latest available statistics on foreign language teaching in the school year 2004/05. Since these data cannot account for changes over the last 15 years, these figures should be understood as rather crude indicators for the current situation of foreign language teaching at Austrian schools, even if the broader developments and their implications remain the same.

offered (Czech, Croatian, Hungarian, Slovakian, Slovenian, Russian and other languages such as Chinese and Japanese) are not even studied by 1% of the learners at upper secondary level (De Cillia & Krumm 2010: 161–162). While these figures can, to an extent, be attributed to the fact that some vocational schools (BMS and BHS) do not offer a large amount of language teaching (De Cillia & Haller 2013: 154–156) because of the wealth of other subjects they have to teach, it appears that the very limited selection is mostly due to a belief that English is the most useful language to learn, as is also indicated by the Eurobarometer survey (European Commission 2012b) described at the beginning of this section.

In addition to regular foreign language teaching, many schools are also involved in CLIL initiatives and offer content lessons taught through a language other than German, which in line with broader developments is usually English. However, for reasons of scope such specialized forms of language teaching cannot be discussed here and the interested reader is referred to section 3 of the latest Eurydice report on foreign language teaching (European Commission/EACEA/Eurydice 2017) for a general overview and to recent publications by Austrian CLIL researchers (e.g. Dalton-Puffer et al. 2018, Dalton-Puffer & Smit 2016, Smit & Finker 2018) for studies conducted at different school types.

Zooming in on academic upper secondary schools (AHS), the figures concerning the foreign languages taught are similar to the wider setting. The Austrian Language Education Policy Profile (Bundesministerium für Unterricht, Kunst und Kultur, Bundesministerium für Wissenschaft und Forschung & Österreichisches Sprachen-Kompetenz-Zentrum 2007: 150, Table 21) shows that in the school year 2004/05 99% of learners in grade 10 studied English followed by French (56.5%), Italian (23.6%), Spanish (16.2%) and Russian (2.3%). Again, English is the predominant foreign language and since it is also the subject of the empirical study presented, further information about English teaching at AHS is a useful background for the interpretation of the results presented in Chapters 6 and 7.

As the first foreign language English is taught for 4 hours a week during grades 1 and 2, and 3 hours a week from grade 3 to grade 12, although individual schools may change these provisions in the curriculum (Bundesministerium für Unterricht und kulturelle Angelegenheiten 2000) to some extent based on the Austrian principle of school autonomy. Taking into account that the average Austrian school year has 39 weeks, this means that according to standard curriculum regulations learners will have had approximately 546 hours of English instruction at the end of lower academic secondary school and just below 1000 at the end of upper secondary school.[76] Similar to all other Austrian school types, the teaching approach for foreign languages specified in the AHS curriculum follows communicative principles and emphasizes the development of social and intercultural competences as well as the acquisition of strategies for future autonomous learning (Bundesministerium für Unterricht und kulturelle Angelegenheiten 2000). For upper secondary level the principle of action-oriented language competence is added,

---

[76] In practice, this estimate is likely to be slightly lower due to holidays within school weeks and lesson cancellations due to illness and other causes.

according to which students should be able to use the foreign language effectively and appropriately in a wide range of situations in private, professional and public contexts (Bundesministerium für Bildung, Wissenschaft und Kultur 2004). The learning aims are specified in terms of the CEFR: at the end of lower secondary school (grade 8) students should have reached level A2 in all four skills with specific B1 competences added for listening, reading and writing (Bundesministerium für Unterricht und kulturelle Angelegenheiten 2000) and at completion of their secondary education in grade 12 they should have achieved level B2 across all four skills (Bundesministerium für Bildung, Wissenschaft und Kultur 2004).

Data on the fulfilment of these learning goals are provided by the results of educational standards tests (*Bildungsstandards*) for grade 8 and the standardized school leaving exam for grade 12. Results of the last educational standard testing for English (Schreiner & Breit 2014) show that 86% of all AHS students tested (N = 26,076) reached level B1 or a higher CEFR level in grade 8, whereas that was only true for 30% at new middle schools. Overall, however, more than 90% of all students in lower secondary schools reached the curricular aim of CEFR level A2. The latest available A-levels results for the school year 2018/19 (Bundesministerium für Bildung, Wissenschaft und Forschung 2019), which is the year in which participants of this study took their A-levels, present a similar picture: at academic upper secondary schools only 8.4% of the learners failed the standardized exam, which uses CEFR level B2 as its pass level.[77] Moreover, 52.4% of all AHS students received the two best grades *Gut* [Good] and *Sehr gut* [Very good], which indicates that the majority of young Austrians attending academic upper secondary schools surpass the required level of English by far.

In sum, this section shows that, next to German as the language of schooling, English is the dominating language across all school types (De Cillia & Krumm 2010: 162) with almost 99% of learners studying it (Nagel et al. 2012: 86). This seems to be a wider European phenomenon, as 97.3% of all students in lower secondary education across the EU studied English in 2014 (European Commission/EACEA/Eurydice 2017: 13). Austria is therefore comparable to other EU countries, except for the fact that at lower secondary level only one foreign language is mandatory (European Commission/EACEA/Eurydice 2017: 12), which is in stark contrast to the EU language policy of L1+2 and the early teaching of two foreign languages recommended by both the European Commission (1995: 44) and the Council of Europe (2002: 19). Thus, even though education is the only context in Austria for which explicit language policies exist, these are in opposition to the EU regulations. As a consequence, Austria is far from reaching the wider policy goal of L1+2 during compulsory schooling (De Cillia & Krumm 2010: 166).

4.2.2 English in the public sphere

Little research is currently available on the role of English in the Austrian public sphere; therefore, the following section will exemplify the uses and functions it fulfils by presenting

---

[77] These students have the option to take an additional oral exam as a compensation, if the students who passed this compensatory exam are taken into account only 1.8% failed their A-levels in English

studies on the physical environment of the linguistic landscape, the position of English in Austrian businesses, and access to English-language media.

The linguistic landscape (LL), which includes '[t]he language of public road signs, advertising billboards, street names, place names, commercial shop signs, and public signs on government buildings' (Landry & Bourhis 1997: 25) and more, is the most obvious instantiation of English in the public sphere. So far, research on linguistic landscapes has mostly been conducted in the Austrian capital Vienna with the largest project to date 'English in the linguistic landscape of Vienna, Austria (ELLViA)' (Soukup 2016) currently underway; hence, results are not yet available. Previous smaller-scale studies do, however, provide first insights into the extent of English use in the Viennese LL. Kral (2012) and Podrepschek (2016) both researched a highly multicultural and multilingual market area, in which 29 languages were identified (Podrepschek 2016). Kral (2012), who also conducted interviews, found that German is used as a lingua franca at the market, whereas in an adjacent square languages such as English, Italian and French, which are not frequently spoken by inhabitants of the area, are used to advertise cafés and pubs. Similarly, Podrepschek's (2016) quantitative analysis of the same area shows that German is the most frequently-used language (65%) followed by English (23%), Turkish (6%) and Bosnian/Croatian/Serbian (3%) (Podrepschek 2016). Her results corroborate Kral's since English mainly appears on transgressive signs, such as stickers or graffiti, as well as commercial signs, where English loanwords or brand names appear to be used to evoke associations with internationality and prestige.

Piritidis (2014) carried out a similar small-scale project in a shopping street and found that while German was the predominant language on monolingual signs (81%), a combination of German and English was favoured on nearly 90% of the multilingual signs. Thus, both Piritidis's (2014) and Podrepschek's (2016) studies show that next to German English is the dominating language in the linguistic landscape of Vienna. The use of English is however not limited to the capital, as might be assumed due to the more urbane and international population. Schlick (2002, 2003) provides examples from Klagenfurt, the capital of Carinthia, as well as a smaller provincial town in Styria and shows that between 20% and 36% of all shop signs contain an English element. In sum, first results on the use of English in the physical environment of the public sphere suggest that English is the most important language next to German and is seen to have considerable prestige. Its influence is however not limited to the linguistic landscape, as the next example from a very different domain shows.

Nowadays, knowledge of foreign languages in general and English in particular is a key competence for businesses and their employees (Schöpper-Grabe 2009: 150). Austrian businesses are no exception; Nagel et al. (2012: 89) report that Business English is widely practised and results of the few empirical studies show that, as in other contexts, English plays an exceptional role in Austrian companies due to its function as a global lingua franca. One of the largest studies on the use of and need for foreign languages in Austria is Archan and Dornmayr's (2006) online survey, which reached 2,017 Austrian businesses in all nine federal

states including small and medium-sized enterprises as well as larger companies with more than 250 employees (see also Tritscher-Archan 2008). Results show that for 12% of these businesses German is not the main company language with most of them using English instead.[78] In addition, 86% of the companies reported needing foreign languages and 80% see English as essential. Interestingly, 45% of the participating businesses stated that English is needed by all employees, whereas in the remaining companies it was mainly the managers who need foreign language skills. The survey further showed that staff mainly use foreign languages for oral communication (70%); in addition, about 60% of the participating companies indicated that they are frequently needed for reading and just over 50% also use them for writing.

This study is complemented by a smaller survey of 40 large companies with a total of 261,677 employees, which were included in a list of the most important Austrian businesses (Weber 2008). For the majority of these companies foreign language skills play a very important role and Weber (2008) reports that English is evaluated as the most important foreign language by far, followed by Hungarian, French, Italian, Czech, Russian, Spanish, Polish and other languages. Foreign language skills, and especially English skills, are regarded as important for staff at all levels in these large corporations, but as particularly crucial for managers. Activities carried out in a foreign language include correspondence and phone calls, reading technical literature, some translations and, mostly at management level, presentations and negotiations. The need for English is particularly prominent in 24% of the 40 companies, which use it as their language of business and in another 24%, where English is used as company language besides German and/or another language. Hence, in addition to English being perceived as the most important foreign language by the vast majority of the diverse businesses in Archan and Dornmayr's (2006) study, its use as a company language in nearly half of the 40 large Austrian enterprises investigated by Weber (2008) highlights the role English has come to play for the Austrian economy.

In contrast to the two examples above, the Austrian media landscape presents a very different picture at first glance. Both print media as well as TV and radio broadcasts are monolingual German, meaning that all international film productions, both for cinema and TV, are dubbed (Busch & Peissl 2003: 190, Nagel et al. 2012: 87). Dubbing is typical of all German-speaking countries in Europe with Germany in the leading position due to its large population. This means that via TV Austrians are mainly exposed to Austrian German and German German, which is frequently used in dubbed versions (Busch & Peissl 2003: 191) and on the many private German TV channels. In fact, data by the Austrian Broadcasting Corporation (*Österreichischer Rundfunk*, ORF) show that Austrian households receive 104 TV channels on average, 83 of which are German, which leads them to the conclusion that Austrians have access to one of the most extensive ranges of programmes in their national language within Europe (Österreichischer Rundfunk n.d., Household equipment and reception). The low linguistic diversity of Austrian

---

[78] In 10% of the businesses English is used exclusively as company language, whereas 2% use English in combination with another language (Archan 2006: 47).

media has, however, been criticized by linguists such as Busch and Peissl (2003) or Purkarthofer (2013) in light of the influential role of media in terms of implicit and explicit language policy, as the use of linguistic resources in the media can shape the linguistic environment of a state or region and is often regarded as a type of standard (Purkarthofer 2013: 242).

In addition, the predominance of German in the Austrian media presents a major difference between my research context and that of previous EE studies, which were predominantly carried out in so-called subtitling countries (see section 2.3). This difference is highlighted by data from the Eurobarometer 2012, in which most participants from Sweden (96%), Finland (95%), Denmark (93%), the Netherlands (93%) and Belgium (51%) fully agreed that they prefer watching foreign films and programmes with subtitles rather than dubbed, while only 14% fully agreed in Austria and the majority of 41% strongly disagreed (European Commission 2012b: 118–119). Because of the largely monolingual media landscape using films, television and radio only was the sixth most frequent foreign language activity among Austrians in 2012. Only 22% stated that they regularly use foreign languages for audiovisual media, which was far below the EU average of 37% (European Commission 2012a).

While it may seem that English is irrelevant in the Austrian media landscape, specific examples as well as recent developments indicate that this conclusion would be erroneous because among the minority, migrant and foreign languages used English again holds a special position. The Austrian Broadcasting Corporation (ORF), which is "the largest media provider operating four national television and twelve radio channels" (Österreichischer Rundfunk n.d.), has to provide an 'appropriate' share of the programme in the languages of the autochthonous minorities (Bundesgesetz über den Österreichischen Rundfunk: §4, section 5a) and one nationwide radio channel with a mainly foreign language programme (Bundesgesetz über den Österreichischen Rundfunk: §5, section 3). Apart from the specific programmes for minority languages prescribed by law (for more information see Busch & Peissl 2003, Purkarthofer 2013), English is the default language used in the few cases where a language other than German is employed in mainstream media: the legally required foreign language radio station is implemented through radio FM4, which broadcasts primarily in English with intermittent use of German (Österreichischer Rundfunk n.d., Information on the programme of FM4).[79] In addition, the few dual-language TV programmes, for which users can choose between audio channels in two languages, always provide a default German and an optional English option.

Furthermore, recent developments have contributed to the much wider availability of English-language media in Austria. While in previous decades foreign language enthusiasts had to buy DVDs or Blu-rays to enjoy films and series in a language other than German, the rise of the internet and internet-based media consumption has made access to such media much easier. In general, Austrian households are well-equipped with media devices (Österreichischer Rundfunk n.d., Household equipment and reception): 87% of all Austrians above the age of 14 have at least

---

[79] In addition, radio FM4 also broadcasts the news in French (Österreichischer Rundfunk n.d., Information on the programme of FM4).

one computer (including tablets, desktops, laptops) at home and the number of those who have a laptop or tablet at their disposal is increasing (73% and 42% respectively in 2016). Similarly, 85% have internet access at home and 96% of all Austrians above the age of 14 possess their own phone, 63% of which are web-enabled smartphones. Data on the use of these devices are provided by the Austrian Internet Monitor, which surveys 1,000 Austrians above the age of 14 four times a year via telephone interviews. The 2017 findings show that 86% of participants use the internet and 71% of these users report going online almost daily. Most internet users are between 14 and 49 years old with 100% of all 14- to 19-year-olds surveyed reporting that they use the internet (Integral Markt- und Meinungsforschung 2017b). To access the internet, the majority use their smartphones (68%)[80] followed by laptops (58%), desktop PCs (44%), tablets (34%), smart TVs (23%) and gaming consoles (20%) (Integral Markt- und Meinungsforschung 2017a). In terms of time spent online, another survey of over 1,000 Austrians between the ages of 16 and 69 who used the internet at least once in the 3 months before data collection shows that the average time spent online on a week day is 170 minutes on a laptop or PC and 114 minutes on a smartphone (IAB Austria, BDVW & IAB Switzerland 2016). Consequently, in total Austrians spend more than four and a half hours online on an average week day (4 hours 44 minutes), which again highlights the pervasive role of the internet.

The figures reported above indicate three things: first, Austrians are well connected to the internet and use it extensively, making it likely that they encounter languages other than German in online media and other websites. Already in 2012, 32% reported that they regularly use foreign languages on the internet (European Commission 2012a) and over the last years these numbers may have changed considerably as internet access continues to rise. Second, many Austrians access online contents on their mobile phones, opening the door to media usage everywhere and at all times. Third, media devices at home are increasingly connected to the internet, even if that is not their primary function, as is the case with gaming consoles or TV sets. In fact, a survey carried out in 2016 shows that 98% of Austrian households own at least one TV set and 23% of these are smart TVs, which are directly connected to the internet (Österreichischer Rundfunk n.d., Household equipment and reception). This development indicates another recent change, namely the rise of streaming services which allow access to films and series on demand. A side effect of platforms like Netflix, Prime Video or Sky is that content is usually available in the original as well as dubbed versions, thus facilitating access to English-language productions.[81] Despite the fact that classic – and thus German – TV still is the market leader, about a quarter of the Austrians watch audiovisual media online on a daily basis, even if in terms of time only about 10% of the total consumption of audiovisual media are accounted for by internet platforms (Rundfunk und Telekom Regulierungs-GmbH (RTR) 2017:

---

[80] A further 34% use other mobile phones to sometimes access the internet (Integral Markt- und Meinungsforschung 2017b).

[81] The content presented on these platforms is frequently produced in English-speaking contexts, particularly in the USA; hence, the majority of original versions use English, but to an extent streaming services also facilitate access to films and series in other foreign languages.

8). In general, younger Austrians tend to use online video platforms more frequently, in the age group between 14 and 29 about a quarter of the total TV time is spent online. Within the category of online video, YouTube and other free video platforms are used by 35% of all Austrians above the age of 14, fee-based streaming services such as Netflix or Prime Video by 26%, videos on social media by 17% and file sharing and torrent platforms by 9%. Among the 14- to 29-year-olds YouTube is the most prominent platform (37%), followed by Netflix (12%), Prime Video (10%), Facebook (6%) and the file sharing platform Burning series (5%).[82]

These data show that a prematurely drawn conclusion that English is irrelevant in the Austrian media landscape due to predominance of German would definitely be wrong. Indeed, English again takes a special and leading position in comparison to all other foreign, minority or migrant languages because it is the default option when foreign language content is required, as in the case of the radio station FM4, and because English-language media are now much more widely available via streaming services than media in any other language. Consequently, it is likely that Austrians increasingly engage with media content in English, which again positions it as the second most important language next to German, as is also the case in the linguistic landscape of various Austrian cities and towns and in Austrian businesses. Due to its wide use and the special status among languages other than German, it is questionable whether English should still be called a 'foreign' language in the Austrian context (Smit & Schwarz 2020). Similar considerations have been voiced, for instance, by Gnutzmann and Intemann (2005) for Germany and Pfenninger and Singleton (2017) for Switzerland. In fact, Pfenninger and Singleton (2017: 13) state that "Switzerland is on the verge of transitioning from an EFL country to an English as an L2 country by virtue of the fact that English is used an intranational lingua franca in a number of domains". While Austria is not inherently as multilingual as Switzerland with its four official languages, a similar transition is taking place here, too, as English has assumed the role of 'next most important language' in addition to German. This situation can be described as 'globalized bilingualism' (Smit 2004), which "understood as the combination of German as prime language and English as default additional language, has turned out to be a valid descriptor for language practices and policies in Austria" (Smit & Schwarz 2020: 309).

4.3   Adolescence in Austria

Adolescents are defined by the WHO as "young people between the ages of 10 and 19 years" (World Health Organization 2018), but definitions vary in relation to comparable terms, such as youth or young people, and in different contexts. In terms of policy, the Austrian laws concerning the representation and promotion of youth (Bundesgesetz über die Vertretung der Anliegen der Jugend, Bundesgesetz über die Förderung der außerschulischen Jugenderziehung und Jugendarbeit) concern all young Austrians below the age of 30, whereas legally the age of

---

[82] Interestingly, another study found that Austrians are not particularly willing to pay for online content in general (Interactive Advertising Bureau (IAB) Austria 2017): 19% of the participants (1000 in total) would be willing to pay for entertainment content and games, but 57% definitely would not. Younger participants and those with higher education would be more willing to pay for content: 43% of the 14- to 19-year-olds would be willing to pay and 28% among the 20- to 29-year-olds.

18 marks the beginning of adulthood and the end of youth (see for instance the Viennese youth protection law: Gesetz zum Schutz der Jugend). Similarly, varying definitions are employed in surveys and research studies (see also Bundesministerium für Familien und Jugend 2016) ranging from a narrow age range (e.g. 15 to 19 years in Blanke & Cornelißen 2005) to a wider population between 11/12 and 18/19 years in youth media studies (e.g. Feierabend, Plankenhorn & Rathgeb 2017, Education Group GmbH 2017a, Waller et al. 2016) or 14 to 30 years in the report on the situation of youth in Austria (Bundesministerium für Familien und Jugend 2016). In light of the focus on 15- to 16-year-old Viennese teenagers in the empirical study, this section focuses on young people between the ages of 14 and 18 although findings reported sometimes also concern younger or older adolescents. In addition, a comprehensive overview of adolescence in Austria is well beyond the scope of such a short section, as youth studies are an established research field of their own.[83] As a result, only a few pertinent points relating to the teenagers' lifeworlds are highlighted in relation to their leisure time before presenting results of recent surveys on spare time activities and media usage in Austria and other German-speaking countries.

As all stages of life, youth is subject to ongoing transformations in line with wider global developments. Großegger (2017: 9) asserts that social, technological and cultural changes have an enormous impact on adolescence so that young people today have vastly different experiences from their parents in many areas of life. Potential reasons given include a pluralization of life plans, fewer social norms, increased competitiveness, and the changing experience and documentation of everyday life through digital media. In addition, Großegger (2017: 12) argues that young people potentially are pioneers and agents of change because they are quicker to take up new ideas and concepts than the older population, be it in form of commercial products or sociocultural innovations. Du Bois-Reymond and Chisholm (2006) identify three key dimensions of change in young Europeans' lives with the first concerning education, training and the labour market, the second referring to relations between different generations, and the third relating to changes in time and space relations. The latter, which is clearly related to the phenomenon of globalization and the rise of digital media and online communication, is of interest to the present study as virtual communication and new connections across time and space are expected to have a major impact on young Austrian's free time activities and potentially also on their language use. As we have seen in the previous section and in Chapters 2 and 3, within these new (virtual) spaces opening up through changing relations, English occupies a privileged position in comparison to other (foreign) languages. In fact, Du Bois-Reymond (2010: 399) highlights this point when she states that in Europe English has become the lingua franca for youth researchers as well as for young people themselves.

As mentioned above, this study is mainly concerned with leisure time, understood as the free time at one's disposal for entertainment and relaxation once all professional/scholastic and

---

[83] See for instance the recent handbooks by Furlong (2017) or Wyn and Cahill (2015).

domestic commitments have been fulfilled (see Statistik Austria 2009: 86). Leisure time can be further differentiated into *structured leisure*, which includes sports, clubs and other organized activities, and *recreational leisure*, referring, for instance, to media use or social activities (Silbereisen 2003). While structured free time activities and especially sports certainly play a large role in the lives of young people, extramural English is closely connected to recreational activities, which can be carried out individually or in groups, but are generally unstructured. This does, however, not mean that recreational leisure activities are free from constraints; in fact, as Flammer and Schaffner (2003: 75) put it "free time for European adolescents is free only within specific limitations". Spare time activities are related first and foremost to the availability of free time in a given sociocultural context, relating for instance to timing in different school traditions, social and family values, as well as to access to resources. However, following Beck's modernization theory it is assumed in youth research

> that young people's organization of free time is controlled now by socio-cultural traditions and social origins to a lesser degree than it was in the past. Rather, commercialised global fashions provide leisure time provision for young people, who absorb them in a highly subjective manner (Blanke & Cornelißen 2005: 512).

Leisure time thus "is in constant evolution [... as a] result of new innovations and trends" (Auger 2016: 173, see also Thole 2010) and represents a space for individual needs and interests, constituting an autonomous sphere of life for young people that offers a contrast to everything connected to requirements and obligations (Großegger 2014: 7). It should thus have become clear that (recreational) leisure time is a vital area of young people's lives that is subject to contextual constraints and to highly individualized and subjective interests (Thole 2010: 736–745). In the following, we will explore the amount of free time as well as the resources available to young Austrians before turning to recent findings regarding spare time activities and media usage.

Previous studies have found that "the large majority of European adolescents have a sizable amount of daily free time, much of it spent on electronic media use, peer socialization, and sports" (Flammer & Schaffner 2003: 75). Investigating German adolescents' time use between 1991 and 2001, Blanke and Cornelißen (2005) found that the amount of free time remained essentially unchanged with an average amount of 5.5 hours on school days and 9 hours on weekends (see also Thole 2010: 738). Zuzanek (2005) conducted a cross-national comparison between 1980 and 2003 and found that, despite some culture-specific differences, overall there were striking similarities concerning young people's lifestyles in developed industrial societies:

> A comparison of adolescent time use and its changes over the past twenty years reveals strong convergent trends. In most countries, adolescents gained additional free time. Mass media consumption and the use of electronic media expanded. On school days, teens went to bed later, but were compensating for this by longer sleep on Sundays. Reading and eating at home declined. Adolescents' life-style preferences and contents of their free time activities [...] also show considerable cross-national similarities (Zuzanek 2005: 396).

Therefore, it is to be expected that Austrian adolescents' lives are relatively similar to those of their European or North American counterparts. The latest data on the amount of free time

available to Austrian teenagers stem from the last time use survey conducted in 2008/2009 (Statistik Austria 2009). According to these findings adolescents between 10 and 19 years of age have an average amount of 4 hours 17 minutes of free time per day (Monday to Sunday, see Statistik Austria 2010). Newer time use data from Germany, which is comparable to Austria in many respects, collected in 2012/13 (Statistisches Bundesamt 2015a) show that 10- to 17-year-old Germans had 6 hours and 38 minutes of leisure time (Statistisches Bundesamt 2015b). The sizable difference between the German and Austrian estimate can at least partly be explained by different definitions of leisure time: while the Austrian time use survey apparently did not count social contacts as free time, the German survey did. Therefore, it seems reasonable to add another hour and 41 minutes for social contacts and volunteering to the Austrian estimate, resulting in a more plausible and more comparable result of 5 hours and 58 minutes of leisure time per day for young Austrians (see Statistik Austria 2010).

Next, the question is which resources Austrian adolescents can make use of during their spare time. One response is provided by the Upper Austrian youth media study (e.g. Education Group GmbH 2017a), which has surveyed 500 teenagers between the ages of 11 and 18 and their parents and teachers in the region of Upper Austria five times since 2008. Since there is no nationwide youth media study, comparable to the German JIM study (e.g. Feierabend, Plankenhorn & Rathgeb 2017) or the Swiss JAMES study (e.g. Waller et al. 2016), these data have to be used as an approximation for the whole country. Even though Upper Austria is more rural than Vienna, the results are indicative of the wider Austrian situation as many activities are carried out online and are thus not bound to a particular geographical location. In addition, the authors of the study also argue that the data can be used as a reference point for the whole country (Education Group GmbH 2017a: 3). Hence, in the following the Upper Austrian data will be described to provide information on young people's leisure time preferences, resources and activities, and they will be compared to the German and Swiss data to show that these developments are part of a larger trend in the German-speaking part of Europe and probably also in other geographical areas.

In terms of resources (Upper) Austrian teenagers are rather privileged as they have access to a wealth of different media at home (compare also the more general data reported in section 4.2.2). Almost all of the 500 adolescents surveyed have access to a TV (96%), a computer or laptop (94%) and the internet (94%). In terms of other media equipment more than three quarters of the households also own a radio (85%), a stereo or CD player (78%) and a DVD player (79%), and more than half of the teenagers have a tablet (65%), an mp3 player (60%), and a fixed (60%) or portable (56%) gaming console at their disposal at home. Interestingly, 64% of the young Upper Austrians surveyed also have access to a daily newspaper. Concerning devices owned by the adolescents themselves, smartphones are clearly in the lead (85%) followed by computers or laptops (59%), music devices such as radios (52%), stereos (48%) or mp3 players (41%), and gaming consoles (39% each for portable and fixed consoles). However, only 36% have their own TV set; there is a marked decline in teenagers' possession of TVs from 51% in 2008 to 36% in

2017, which suggests that traditional TV may be losing importance in this age group (Education Group GmbH 2017c: charts 6, 9).[84] Considering that relatively young participants from the age of 11 onwards are included in these results, they indicate broad access to different media. In fact, the figures, which are unfortunately not available split by age groups, may be considerably higher for teenagers over the age of 14 as the youth trend study *TRacts* (cited in Bundesministerium für Familien und Jugend 2016: 63) showed that already in 2014 96.7% of the 14- to 19-year-olds reported owning a smartphone, which is a considerably larger proportion than found in the 2017 Upper Austrian youth media study for 11- to 18-year-olds.

In comparison to Austria, German and Swiss data generally show a slightly higher media saturation in the households of the 12- to 19-year-old participants: among the 1,200 adolescents surveyed in Germany 97% possess their own smartphone, 69% a computer or laptop and 53% a TV set (Feierabend, Plankenhorn & Rathgeb 2017: 6–8) and of the almost 1,100 participants in Switzerland 99% own a smartphone, 76% a computer or laptop and 53% an mp3 player, but only 30% have a TV set (Waller et al. 2016: 13–17). In general, the figures for access to further devices in adolescents' homes are also slightly higher than in Austria, which to an extent may be due to the larger sample sizes of the JIM and JAMES studies and to their slightly older age range. An interesting insight from the JIM study is that in Germany ownership of media devices increases with age: 15/16-year-olds are more likely to have their own laptop, computer, gaming console or TV set than younger adolescents, supporting the conclusion on age effects drawn above. Only smartphones are owned by almost all teenagers from the age of 14 onwards (Feierabend, Plankenhorn & Rathgeb 2017: 9–10). Furthermore, the German data show that 54% of teenagers have access to streaming platforms like Netflix or Prime Video at home, while daily newspapers are only available in 48% of the households (Feierabend, Plankenhorn & Rathgeb 2017: 6–7). In Switzerland the situation is slightly different: 38% have a subscription to a streaming service for films and series and 29% for music, whereas subscriptions to daily newspapers are still held by 59% of the participants' families (Waller et al. 2016: 15). While the proportion of subscriptions to newspapers is even higher in Austria (see above), no data were collected on the availability of streaming services among Upper Austrian teens, but it is likely that the number of households with subscriptions is rising, as is the case in Germany and Switzerland. Having established that, similar to their German and Swiss counterparts, Austrian adolescents have ample access to different media in their homes and are likely to own at least a smartphone themselves and potentially also a computer or laptop, the next relevant aspect is how they use these resources and which other activities they like to engage in.

In general, a large part of young Austrians' free time is governed by friendships: meeting friends and communicating with them in real and digital environments are the two top leisure activities. In addition, the intensive use of computers, tablets and especially smartphones has become a

---

[84] These figures were taken from charts summarizing the results of the survey among 500 Upper Austrian 11- to 18-year-olds (Education Group GmbH 2017c); in comparison, the results of the survey among 207 parents (Education Group GmbH 2017b) show slightly higher figures for media possession, which may be due to the difference in sample size or to a more comprehensive provision of details by the adults.

natural part of adolescents' lifeworlds (Education Group GmbH 2017a: 4), but they also frequently engage in other pastimes. The top five non-media activities are meeting friends (80%), resting and relaxing (66%), family activities (65%), doing sports (56%) and taking care of pets (47%) (Education Group GmbH 2017c: chart 3). Similar results were found in Germany and Switzerland, which shows that although media have become an integral part of young people's leisure (see also Bundesministerium für Familien und Jugend 2016: 63), they have not displaced other activities. [85] Turning to media use during young Austrian's spare time, the most common media activities overall are communication via social media and messaging services (80%), 'doing something' on a computer, table or smartphone (70%),[86] watching DVDs or YouTube (69%), watching TV (65%), surfing the internet (63%) and playing games on various media devices (57%) (Education Group GmbH 2017c: chart 3).[87] Surprisingly, music-related activities are not found in the top results here, which is likely due to the wording of the response option presented in chart 3: only 47% of the participants report listening to mp3s or CDs, but a more specific interview question shows that YouTube is the most popular source of music (69% use it very frequently), followed by radio (36%), mp3s (35%) and music streaming (28%) (Education Group GmbH 2017c: chart 13).

In contrast to the Upper Austrian data (see footnote 87), the JIM and JAMES studies provide information on the frequency of use for media activities. 97% of German teenagers use the internet as well as their smartphone at least several times a week, with 89% and 93% respectively reporting daily use. Furthermore, 95% listen to music at least on a weekly basis, 86% watch online videos, 75% watch TV, 73% listen to the radio (including online programmes) and 62% play digital games. Other activities such as reading books or using streaming services for music and films or series are used by less than half of the participants on a weekly basis (Feierabend, Plankenhorn & Rathgeb 2017: 13–14). Similarly, the vast majority of Swiss teenagers report using their smartphone (95%) and the internet (85%) on a daily basis. The next most common activities which are carried out at least a few times a week are listening to music (94%) and watching TV (73%); which here seems to include both traditional and online offers, and listening to the radio (50%). Further activities, such as taking photographs, playing digital games, using a tablet or reading (free) newspapers and books are done by less than 50% of the Swiss participants (Waller et al. 2016: 22–23). An activity that is conspicuously absent from the top activities among adolescents in Switzerland is watching video clips; it is not clear whether this was not given as a response option or whether the term 'watching TV' includes all forms of audiovisual media in this survey. Together with the unusually low figures for music in the

---

[85] In Germany the top five activities are meeting friends (73%), doing sports (68%), family activities (34%), making music (24%) and sport events (15%) (Feierabend, Plankenhorn & Rathgeb 2017: 12), while in Switzerland they include meeting friends (76%), doing sports (66%), resting and relaxing (58%), engaging with pets (41%) and making music (23%) (Waller et al. 2016: 10). While some activities are not included in the top five activities across countries, overall the results are remarkably similar.

[86] This answer option appears to have been kept deliberately open as further questions provide more detailed information on activities carried out on computers and online (see Education Group GmbH 2017c: charts 28, 44).

[87] Unfortunately, these data do not allow conclusions on the frequency of use as the question put to participants was with which of the activities shown they spend their free time rather than how often they engage in the activities listed.

Austrian data (see above), this is a good example of the influence of methodology on findings and highlights the difficulty of comparing studies conducted in diverse contexts and using different operationalizations. In general, however all three studies clearly indicate that online activities, which are also frequently carried out on mobile phones, occupy a central position among young people's media activities followed by music, TV, video clips and audiovisual content, and, to a lesser extent, gaming.
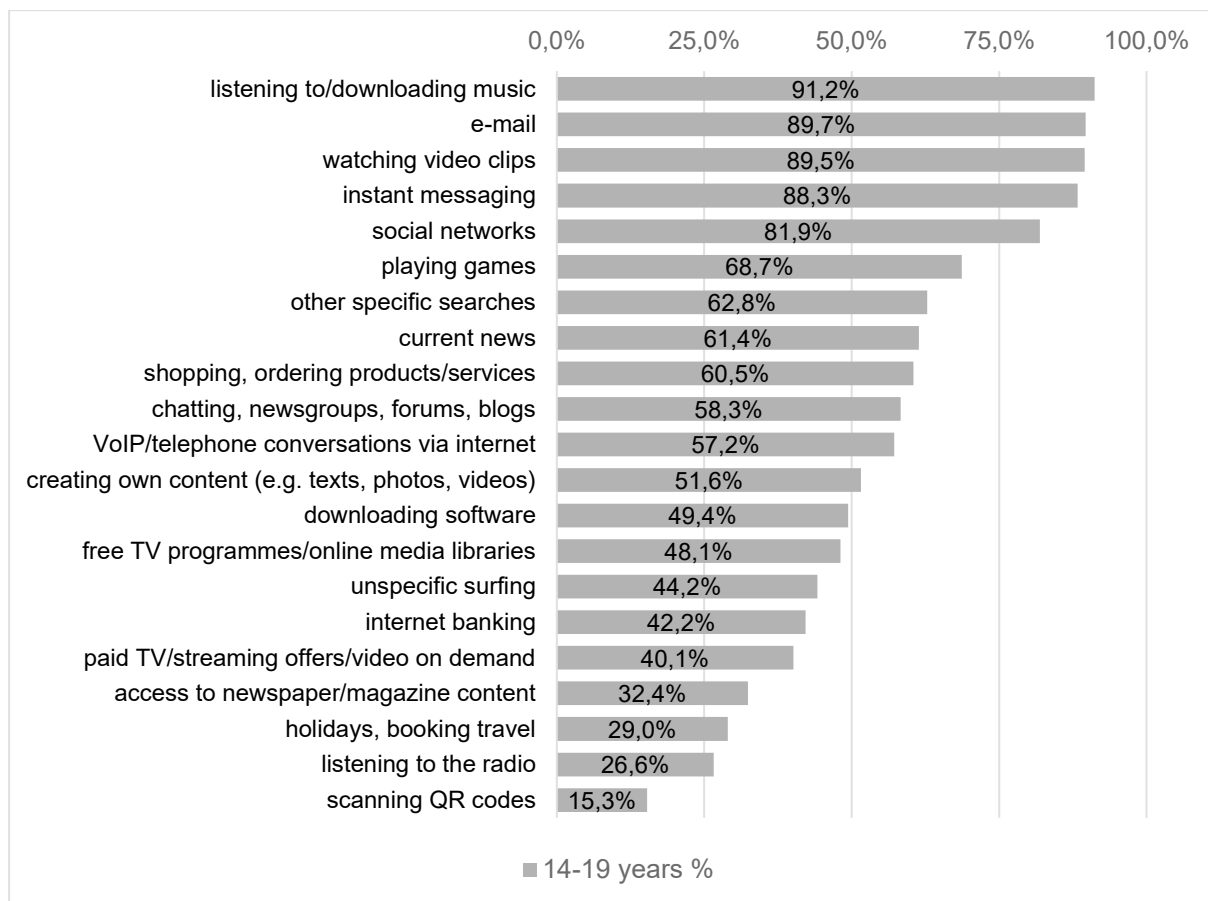
Figure 4.2: Online activities of 14- to 19-year-old Austrians based on a survey of 1,136 participants in this age group in 2017 by Verein Arbeitsgemeinschaft Media-Analysen (2017).

For Austria, more detailed data on online activities are provided by the Austrian media analysis 2017 (Verein Arbeitsgemeinschaft Media-Analysen 2017), which is available for different age groups. It shows that the large majority of 14- to 19-year-old Austrians (*n* = 1,136) accessed the internet on seven days in the week before data collection (92.6%), whereas in the overall sample including all age groups (*N* = 15,562) the proportion was only 66.9%. Moreover, 97.5% of the adolescents were online the day before, more than in any other age group, which underlines the importance of virtual environments for young people. Figure 4.2, which is a summary of data from the Austrian media analysis, displays the purposes of internet use for the 14- to 19-year-olds within the four weeks prior to data collection. As can be seen, music, video clips, and online communication in form of e-mails, instant messaging and social networks are the most popular activities. Young people engage in these activities far more frequently than older Austrians: with the exception of e-mails, there is a difference of a least 30% between the youngest age group and

the overall sample. Teenagers are also much more likely to play computer games, use chats, forums or blogs, watch video content on streaming platforms and create their own content. In fact, the only activities which are carried out more frequently by older Austrians are internet banking, booking travel, and reading newspapers or magazines online (see Verein Arbeitsgemeinschaft Media-Analysen 2017).

The Upper Austrian data shows a similar picture as the most frequent activities carried out (almost) every day on the computer are surfing the internet (58%), using messaging apps (46%), listening to music (46%), and using online communities (39%) and streaming services (29%) (Education Group GmbH 2017c: chart 28). While computers allow the possibility to carry out tasks offline, teenagers spend a considerable amount of time online: on average participants in the Upper Austrian youth media survey who have access to the internet spend 94 minutes (1 hour and 34 minutes) online per day. In line with results of the German JIM study (Feierabend, Plankenhorn & Rathgeb 2017: 30–31) which show that the amount of time spent online increases with age, it can be assumed that older teenagers in Austria also spend more than the average 94 minutes online since this estimate also includes 11-year-olds.

In terms of popularity YouTube (84%) is the number one online activity, followed by searching for information in general (62%) and for school (58%), watching films or videos (62%) and instant messaging (57%) (Education Group GmbH 2017c: chart 44). YouTube is the leading platform for watching video clips (Education Group GmbH 2017c: chart 16) and listening to music, as has been mentioned above, and Upper Austrian teenagers' favourite website in general (Education Group GmbH 2017c: chart 50). The Austrian youth internet monitor (Saferinternet.at 2018) also shows that after WhatsApp with 93%, YouTube is the next most popular app with 90%, followed by Instagram (68%), Snapchat (65%) and Facebook (48%), the use of which has however sharply decreased. In terms of content, music videos, funny clips, entertainment clips by YouTubers, tutorials, videos on fashion and beauty, and let's plays, clips which allow viewers to watch a gamer play a game with running commentary, are most popular (Education Group GmbH 2017c: chart 19, see also a recent study by the Institute for Research on Youth Culture 2018). On average, teenagers spend 49 min on YouTube per day, nearly as much as with traditional TV (56 minutes) and more than on streaming platforms (31 minutes) (Education Group GmbH 2017c: chart 14). The example of YouTube also shows how difficult it is to differentiate between EE activities in online contexts: the platform is used for watching videos for entertainment as well as for information purposes and for listening to music; in addition, it also allows users to actively contribute to the community through comments or by uploading their own videos.

In sum, media are an essential part of young people's leisure time and especially online environments play a crucial and pervasive role as internet access is rapidly becoming ubiquitous in Austria. It is therefore not unexpected that the one device young people could not do without is their smartphone, which has clearly become indispensable. In line with such developments, the internet is now seen less as a topic of interest, but rather as an integral part

of everyday life (Education Group GmbH 2017a: 4). Furthermore, traditional media such as TV and radio are losing ground: as shown above, adolescents spend almost the same amount of time with YouTube clips as with traditional TV programmes. Since more and more teenagers have access to streaming services and video on demand platforms at home, viewing habits will likely continue to change. Potentially, these developments could have a significant impact on young Austrians' language input via (online) media as the currently largely monolingual media landscape (see section 4.2.2) is opening up to other languages and particularly to English as a global lingua franca.

4.4 Extramural English in Austria: the research context of the present study

While the role of English in education and the public sphere has been discussed in section 4.2, this section focuses on extramural English in Austria, which in a way can be regarded as a glimpse at English in the private sphere (see also Smit & Schwarz 2020). Having explored the concept of EE in Chapter 2, previous research on out-of-school English learning in the Austrian context is reviewed here before turning to a close description of the research area investigated by the empirical study presented in the following chapters.

So far, few studies on informal English learning through leisure time pursuits are available in Austria. First insights into learners' EE practices at different educational levels are provided by five studies: a BA paper by Wieland (2016) is the first to examine EE in an Austrian primary school context, two MA theses by Ringl (2014) and Hahn (2017) explore learners' practices in different school types at secondary level, and Miglbauer (2017), another MA thesis, as well as Trinder (2017) investigate informal English learning at tertiary level.

While studies on engagement with EE at primary level (e.g. De Wilde & Eyckmans 2017; Jóhannsdóttir 2017; Kuppens 2010; Lefever 2010; Lindgren & Muñoz 2013; Persson & Prins 2012) are becoming more prominent, in the Austrian context Wieland (2016), who conducted an exploratory case study with one 4th-grade primary school class in Salzburg, presents the only source of data at present. For the study, 23 9- to 11-year-old children recorded their out-of-school English activities in a structured language diary at the beginning of 13 school days. Information on the amount of time spent with English outside school on the previous day was gathered in six major categories: (1) listening to music and singing, (2) watching films, series, video clips and advertisements, (3) reading or being read to from books, newspapers and magazines, (4) playing digital games on a computer or phone, (5) surfing the internet, and (6) speaking English to family member or friends.

On average the children spend 21 minutes with English per day, but the overall amount of English input varies strongly with one boy reporting no exposure to English and four pupils less than one hour over the entire data collection period, whereas seven of their classmates spent more than 6 hours with English overall. In terms of preferred EE activities, music was the most common activity with which the pupils spent more than half of the total EE time (55.7%),

followed by games (21.6%), films/series/videos/advertising (12.7%) and conversations in English (8.6%). The predominance of music as a source of English input is to be expected (see section 3.3.1) and it is plausible that gaming and watching TV – even if in English – are popular in this age group. What is, however, quite astonishing is the relatively large amount of time spent speaking English (8 hours in total). Wieland (2016: 43) attributes this result to the influence of two girls in this class, who regularly attended ballet classes with an English-speaking instructor. Contrary to her hypotheses, she did not find a relation between elementary school pupils' engagement with EE and their motivation and language competence; however, this finding may be influenced by the methods used: motivation was measured using one item only and information on pupils' language competence was gathered through an expert interview with their class teacher. Hence, the instruments are likely not sensitive enough to show any effects and more extensive studies are needed to investigate the impact of EE on language development in this age group. Still, Wieland (2016) succeeds in showing that already very young Austrian learners of English between 9 and 11 years of age come in contact with English outside school, although there is great variation in terms of amount of exposure.

While the findings of this small-scale study provide interesting first insights into the EE environments of primary school pupils, the majority of available studies have focused on older learners. Ringl (2014) constitutes a first attempt to investigate the use English in out-of-school contexts among upper secondary school pupils. 169 students attending their penultimate school year in a general academic school (AHS, grade 11) in Vienna and two vocational colleges (BHS, grade 12) in rural Lower Austria participated in a survey on their use of English outside school. Results indicate that the students come in contact with English mostly through media. About 12% of participants report watching English films, 14% watch TV in English and 23% play games several times a week.[88] The more urban Viennese sample reports a more extensive use of English media than the Lower Austrian students, but the samples are not directly comparable as they differ in terms of school type and other crucial factors such as socioeconomic background or degree of multilingualism.

A more detailed and theoretically well-grounded investigation is Hahn's (2017, 2018) study, which has already been summarized in section 3.3.1. Together with the present study, Hahn's project is the first to systematically explore the relationship between EE and language learning, more specifically vocabulary acquisition, in Austria. The EE practices of Viennese students attending vocational business middle schools, who most frequently encounter English outside school through music, audiovisual media and social media, show significant positive relations with both receptive and productive vocabulary size and thus indicate that EE activities can have an impact of language development. A comparison of Hahn's promising findings with those of the present study will certainly yield interesting insights and will be discussed in Chapter 8.

---

[88] Percentages were calculated from results reported in Ringl (2014: 70–78) to make the findings more comparable to the other studies reported.

Moving on to the tertiary level, Miglbauer's (2017) research involves a dual focus on fostering language skills and digital competences in tertiary education through students' EE practices in online and offline contexts. In total, 333 first-year students studying various subjects at two Austrian universities filled in questionnaires in the winter terms between 2013 and 2016.[89] In her analysis, Miglbauer found that four activities were carried out at least once a week by at least half of her participants. These are listening to music (more than 95%), reading webpages (more than 80%), watching films without subtitles (more than 55%) and reading e-mails (more than 50%). With regard to the four language skills, English is mainly used for listening and to a lesser extent for reading, whereas writing and speaking occur much less frequently during students' online and offline leisure activities (Miglbauer 2017: 45–47). Thus, EE activities in general, but especially the most popular ones, mainly involve receptive language use, as is also the case in Hahn's (2017) study. Concerning the more specific focus on digital aspects, almost 80% of the students indicated that they use English for their online activities, with 10% stating that it is their predominant online language. The most common online activities conducted in English are searching for information (more than 80%), reading the news (more than 70%) and written chat communication (more than 50%) (Miglbauer 2017: 44–45). While the popularity of music and films is not unexpected in light of previous studies, the most common online activities as well as the frequent use of e-mails in this sample indicate the importance of these activities in a university context and potentially also reflect the participants' more advanced age.

In a similar vein, Trinder (2017) is mainly concerned with digital English practices and with their relation to informal and deliberate learning. Following on from Sockett's work on online informal learning of English (e.g. Sockett 2013, 2014, Kusyk & Sockett 2012, Toffoli & Sockett 2013), she conducted a questionnaire survey of 175 students at the Vienna University of Economics, which combined closed quantitative items and open qualitative questions. In addition to frequency of use, the survey also included items on the perceived usefulness of online activities for English learning. Results show that students regularly use English in virtual contexts; online dictionaries are the most commonly employed resource, with 94% reporting daily or frequent use, followed by TV, radio, video clips and series with 73%, informational websites with 71%, social networking sites with 58% and online news sites as well as company websites with 45% each. In terms of usefulness for language development, students found that online dictionaries (74%), TV, radio, video clips and series (67%), films on DVDs or Blu-ray (60%) and online news sites or journals (51%) "helped very much" (Trinder 2017: 404–405). Moreover, Trinder was interested in whether such online activities are used deliberately by learners to improve their English, expecting

> the intentional aspect to gain in importance in informal environments, with students deciding to access resources such as news sites in English even though the equivalent content is available in their native language, simply because it benefits their English (Trinder 2017: 402).

---

[89] Students who specialize in English were excluded from the sample to avoid biasing the results (see Miglbauer 2017: 37–38).

Indeed, 72% of students in her sample indicated that they intentionally use online activities to develop specific aspects of their language competence and named online news site and journals as well as audiovisual media as the two resources most frequently used for deliberate online language practice; even though watching audiovisual media evidently also entails an element of pleasure (Trinder 2017: 407). The author states that while many students began to watch TV, series or video clips in English for entertainment purposes, "there has been a shift towards dual purpose engagement" (Trinder 2017: 407). She argues that the fact that students recognize the positive effects of online English activities and actively choose to do them in English, points to learning becoming "a deliberate, even if usually secondary, aim" (Trinder 2017: 407). While this is a novel and interesting viewpoint, the question of whether informal learning through EE activities is intentional or incidental is certainly open to debate and very likely different for each learner (see also sections 2.1 and 7.3.).

In sum, the five studies summarized above allow first conclusions about EE practices in Austria: first, they suggest that English plays a considerable role in the lives of young Austrians from primary school pupils to students at university. Second, they show that media are the main point of contact with English with some activities such as listening to music and watching audiovisual media being popular across all age groups. Third, the most frequent activities identified in these studies mainly involve the use of English in a receptive way; and fourth, the different findings illustrate that there is ample room for further research. Different research foci have been examined in the different age groups with a focus on motivation and language learning at primary level (Wieland 2016), the link between EE and vocabulary learning at secondary level (Hahn 2017), and digital environments and learning technologies at tertiary level (Miglbauer 2017, Trinder 2017).

Adding to this small existing body of research in the Austrian context, the present study explores EE practices and vocabulary learning in academic secondary schools from a quantitative and qualitative perspective; thus, giving learners a chance to voice their perceptions of the phenomenon of extramural English. Before presenting the research design of the empirical study in Chapter 5, a close description of the research context is given below. Embedded in the wider linguistic, educational and socioeconomic context described in this chapter, the more specific research area can be characterized using Benson's (2011) model of language learning beyond the classroom. As described in section 2.1, Benson's original model includes the four dimensions of location, formality, pedagogy and locus of control. More recently, additional dimensions have been proposed, but while these suggestions are highly interesting in terms of theoretical development (see section 2.1), Benson's original model is perfectly adequate for the purpose of characterizing the research environment in terms of the learning situations investigated.

The adaptation of my graphical representation of Benson's (2011) model (see section 2.1) in Figure 4.3 visualizes my conceptualization of learning situations in relation to extramural English, which has also been addressed in Schwarz (2016: 59). Such situations are primarily

defined by their location, which is extramural as the name suggests, and thus necessarily exclude extracurricular activities, which still take place at schools and are organized by the educational institution. Similarly, any formal teaching and learning directly connected to school is excluded, meaning that tutoring lessons, additional classes, homework, other assignments or test preparation fall outside the area of interest of this study. The learning opportunities under investigation take place during participants' leisure time (see sections 2.2 and 4.3) and are thus typically informal. It is the learner who holds control over their learning, which means that locus of control can range from non-direction to self-direction, depending on the learners' aims for and awareness of the learning process.
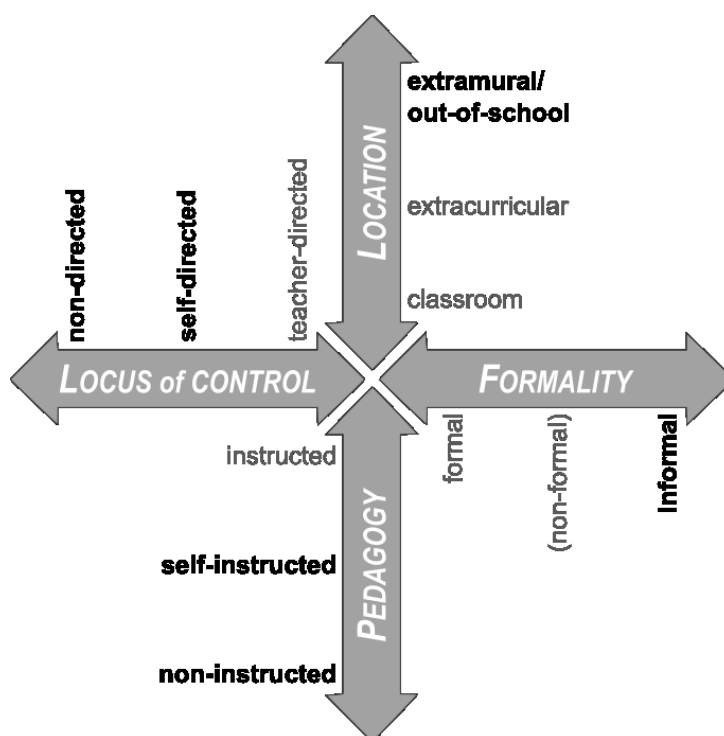
Figure 4.3: Features of the learning situations under investigation in relation to Benson's (2011) framework of language learning beyond the classroom

Likewise depending on the learners' intentions is pedagogy, which can vary between non-instructed, naturalistic settings and self-instructed learning. Hence, a typical learning situation of interest could involve a learner looking up an unknown phrase heard in a YouTube video, which is a self-directed learning opportunity triggered by an informal, non-instructed activity in an extramural context. Similarly, remembering a phrase frequently used in a favourite game through repeated exposure also constitutes an informal, non-directed and non-instructed learning situation characteristic of an EE environment. But a learner, who actively and intentionally searches for specialized vocabulary in relation to photography or any other hobby to understand English texts on this subject also falls within the scope of this study, since this learner engages in self-directed, self-instructed, informal and extramural learning. Hence, as long as an activity involving English takes place in an out-of-school context during young people's leisure time and is (mostly) engaged in voluntarily (see also Sundqvist 2009a: 25–26), it is an EE activity of interest to this study.

4.5   Summary

This chapter has provided information on the linguistic, educational and socioeconomic context of the empirical study presented in the next chapters. We have seen that Austrian language policy does not reflect the linguistic reality, in particular with regard to its diversity, and that apart from a small group of recognized, autochthonous minority languages (Austrian) German is the legal standard. Among the 'foreign' languages, English takes on a special position in Austria: it is the dominant foreign language studied by the vast majority of students across all school levels; next to German, it is the language most commonly found in the linguistic landscape; and it plays a large role in Austrian economy, both for international business and as a company language in some larger corporations. Even in the media landscape, which is dominated by German due to dubbing practices, English is the go-to standard if a foreign language is used. Hence, across different public domains such as education, business and the media there is a linguistic duality with German as the official Austrian standard and English as the default additional language, a situation which Smit (2004) has aptly named "globalized bilingualism". Although many Austrians also speak other languages, English clearly takes on a special position due to its role as a global lingua franca and its perceived usefulness, which highlights that it is currently not a completely 'foreign' language anymore.

Other contextual aspects that are of great importance in relation to the present study concern adolescence in Austria. Austrian teenagers, like their German and Swiss counterparts, have access to a variety of different technological devices and they make ample use of them. Although leisure activities are not free from socioeconomic constraints, the findings of previous media-related research indicate that the Austrian media landscape is changing and potentially opening up to new languages. Hence, the prerequisites for extramural engagement with English, or other languages, in a variety of online and offline contexts are met and first research projects on language use outside formal education indicate that these opportunities are also taken up. The studies by Miglbauer (2017), Ringl (2014), Trinder (2017) and Wieland (2016) show that English plays a role in the lives of young Austrians from primary to tertiary level and that some activities such as listening to music or watching audiovisual media are popular across age groups. However, while these studies focused on digital competencies, learning technologies and motivation, the present study and the approximate replication by Hahn (2017) are the first projects to investigate the relation between out-of-school language use and language development in form of vocabulary acquisition. More detailed information on the specific research aims, questions and design are presented in the next chapter.

5 Research design and methodology

The review of research on extramural English and vocabulary learning from EE activities in Chapters 2 and 3 has shown that there are many questions that remain unanswered at present and that the field of informal (English-)language learning offers many opportunities for research. Chapter 4 presented Austria as a context in which little EE research has been conducted to date. Specifics of media and language contact were highlighted, which distinguish this setting from the subtitling countries in which most European EE research has been carried out. This chapter presents the empirical study conducted in Vienna, Austria and provides detailed information about the aims and questions addressed (5.1) and the mixed methods research design used (5.2), as well as the sampling strategies (5.3.1 and 5.4.1), the development and selection of all instruments (5.3.3 and 5.4.2), the data collection procedures (5.3.4 and 5.4.3) and methods of analysis (5.3.5, 5.3.6 and 5.4.6). Following the structure of the study (see section 5.2.2), the quantitative strand is presented first after an introduction to the overall research design.

5.1 Research aims and questions

In keeping with previous research on extramural English and informal language learning, the first major aim of studying EE in a new context must be to collect data on the extent of engagement with English leisure activities. As Sockett (2014: 32) points out with regard to OILE "the necessary first step is to begin to get to know the online informal learner by measuring how widespread the phenomenon is and what the major areas of the field may be". The same is true when researching the more inclusive concept of extramural English (see section 2.2). Establishing a quantitative overview of learners' practices before studying more specific aspects in greater detail is essential because it allows researchers to "understand whether the phenomena under scrutiny are occasional accidents provoked by a small number of careless Google searches or broader experiences of a significant proportion of users of English today" (Sockett 2014: 148).

Furthermore, quantitative data on EE practices can then be linked to language learning outcomes to explore the relationship between engagement with EE and language development. In the present study the focus is on the relation between EE activities and vocabulary size. The review of existing research in section 3.3.1 shows that several studies have endeavoured to link vocabulary knowledge to informal out-of-school learning. This study contributes to this body of research by investigating the relation between extramural engagement with English and receptive and productive vocabulary size in a new research environment, specifically academic secondary schools in Vienna, Austria. Data from this context are of interest to a wider international audience as well because the Austrian setting resembles other contexts in which English, although used widely in education and business, has no official public role and is not usually employed in traditional media, such as TV broadcasts (see Chapter 4). Therefore, this study complements previous European research, which has tended to focus on so-called

subtitling countries as shown in section 2.3. Furthermore, this study also introduces an innovative methodological idea in relation to productive vocabulary: by comparing the data elicited with the help of the productive vocabulary test Lex30 (see section 5.3.3.3) to the coursebooks used at school, it attempts to establish not only that vocabulary can be acquired outside school, but also investigates which types of vocabulary are potentially learned from EE.

In addition to exploring the connection to language learning outcomes, the ultimate goal of much, if not most, EE research is to find ways in which language learning can be supported across different contexts. Clearly, much more research as well as pedagogical concepts will be needed to achieve this long-term aim (see section 9.3), but a first step must be to gain a better understanding of the learners' views, who are arguably the most important stakeholders in informal language learning. A qualitative exploration of learners' perspectives gives room to their voices as agents in language learning processes and serves to highlight different individual perspectives (see also Sockett 2014: 23). As discussed in section 2.3, to date relatively few studies on EE and informal language learning have included such a qualitative focus on learner perspectives, although it presents a valuable counterpart to quantitative data on EE practices and enhances our understanding of this complex phenomenon.

The aims of the empirical study presented in this thesis are therefore threefold, as briefly addressed in Chapter 1. First, the study aims to map the landscape of extramural English in Vienna, Austria by providing an overview of the English activities Viennese adolescents attending upper secondary school engage in during their leisure time. This goal entails collecting information from a representative number of students and providing information on the amount as well as the type of exposure to English outside school. Second, the study aims to explore the relationship between engagement with EE and receptive and productive vocabulary size, which allows inferences about the potential of EE for vocabulary acquisition. Third, the project also aims to include learners' perspectives. It describes Viennese adolescents' perceptions of the EE phenomenon including their views on the importance of English in their everyday lives, their beliefs about (vocabulary) learning from EE and their thoughts on the link between in- and out-of-school language learning.

In keeping with these three research aims, the overarching research question addressed in this study can be expressed as

> *What is the impact of extramural English on Viennese upper secondary school students' vocabulary knowledge and development?*

This question entails two parts: first, it postulates that Viennese upper secondary school students engage with extramural English and second, it hypothesizes that this engagement affects vocabulary knowledge and development. It is important to highlight that *impact* here is seen as relating both to measurable impact in terms of a quantitative relationship and to perceived qualitative impact in terms of participants' evaluation of the effects of EE on their vocabulary knowledge and learning. Hence, the overarching research question entails both a quantitative and a qualitative perspective and reflects the need for a mixed methods study.

In addition to the overarching research question (RQ), five more specific research questions were formulated taking previous research, the objectives of the study and the practical constraints for the scope of the project into account. These are the five main research questions that guided the design of the mixed methods study:

RQ 1: *What kinds of extramural contacts do Viennese upper secondary school students report having with English?* (mainly quantitative)

RQ 2: *How much time do Viennese upper secondary school students report spending in contact with extramural English?* (mainly quantitative)

RQ 3: *What is the relationship between extramural English and the receptive English vocabulary size of Viennese upper secondary school students?* (mainly quantitative)

RQ 4: *What is the relationship between extramural English and the productive English vocabulary size of Viennese upper secondary school students?* (mainly quantitative)

RQ 5: *What are Viennese upper secondary school students' perceptions of EE and its potential for language learning?* (mainly qualitative)

As can easily be seen, the majority of these question are traditionally associated with quantitative methods and have therefore been labelled quantitative; while the last question is more open and qualitatively oriented.[90] The predominance of more quantitatively-oriented questions is owed to the nature of this study as the first larger investigation of a new group in a new setting and is reflected in the design of the mixed methods study.

5.2 Mixed methods research design

Having introduced the aims and research questions of the empirical study, this section will provide an overview of relevant methodological considerations before describing the design and the procedures used. The methodological background to this project, discussed in section 5.2.1, is formed by mixed methods research. Section 5.2.2 then introduces the overall research design of the empirical study and section 5.2.3 sets out the research steps including all piloting efforts.

5.2.1 Mixed methods research

Originally stemming from the idea that triangulation of different methods can minimize weaknesses inherent in any single method, mixed methods research (MMR) has developed as a third research paradigm since the late 1980s (Creswell & Plano Clark 2011; Dörnyei 2007). Its defining characteristic is an intentional and justified combination of quantitative and qualitative methods, as indicated by the definition in a seminal article by Johnson and Onwuegbuzie (2004: 17): *"Mixed methods research* is formally defined here as *the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language into a single study"* (emphasis in original). [91] Hence, MMR "recognizes, and works with, the fact that the world is not exclusively quantitative or

---

[90] The main qualitative research question was later split into seven sub-questions for operationalization in the interview guide (see section 5.4).

[91] For an overview of different definitions of MMR see, for instance, Creswell & Plano Clark (2011).

qualitative" (Cohen, Manion & Morrison 2011: 22) and thus allows researchers to gain a more coherent picture of any given phenomenon.

MMR has been gaining increasing attention in applied linguistics (Hashemi & Babaii 2013; Riazi 2017) and can be regarded as particularly suitable for applied research projects because it "allows the researcher to explore a research question from multiple angles" (Mackey & Bryfonski 2018: 104). The main strength of MMR thus lies in the fact that a combination of research methods allows different points of view and can lead to stronger inferences based on different data sources. It has also been argued that MMR can be used to answer research questions for which one methodology on its own is insufficient and that it gives opportunity to present a greater diversity of views and voices (e.g. Teddlie & Tashakkori 2003).

Several approaches to classifying MMR designs have been proposed (e.g. Creswell & Plano Clark 2011; Morse 1991, 2003; Teddlie & Tashakkori 2006, 2009). Building on Morse (1991)'s work, these design typologies mostly use the key features of priority and timing to distinguish between different MMR designs. Priority here refers to the predominance of either the quantitative or qualitative strand or an equal balance of the two strands in the overall design, whereas timing *"generally refers to the timing of the collection of the qualitative and quantitative data"* (Creamer 2018: 61, emphasis in original). Based on timing, concurrent and sequential designs (Creswell & Plano Clark 2011) as well as multiphase or multilevel designs involving several iterations can be distinguished (Creamer 2018; Teddlie & Tashakkori 2009). In addition, the degree of mixing in a given research design can also be taken into account (e.g. (Teddlie & Tashakkori 2009).

Despite the growing number of elaborate design typologies, reviews of empirical studies (e.g. Hashemi & Babaii 2013) have identified design variations that do not fit the proposed typologies highlighting the manifold combinatory possibilities offered by MMR. This is one of the reasons why I agree with Creamer (2018) that it is more useful to categorize MMR studies by purpose rather than by design. She argues persuasively that "[i]t is the purpose, not the design, that drives the decision to used mixed methods" (Creamer 2018: 39) and presents an expanded typology of purposes for mixing methods based on earlier work by Greene, Caracelli and Graham (1989). In this influential study, the authors identified five main purposes for MMR after reviewing 57 empirical mixed methods articles: *triangulation*, *complementarity*, *development*, *initiation*, and *expansion*. Greene, Caracelli and Graham (1989) stress that studies can draw one or more of these purposes as a rationale for mixing methods; thus allowing for creative combinations of the five main purposes identified. Creamer (2018: 28–30) adapts and extends this typology and presents six purposes for conducting MMR: *triangulation/confirmatory* to seek confirmation for a single construct, *enhancement/complementarity* to investigate different aspects of a complex construct, *development* in relation to sampling or instrument development, *initiation* to explore different aspects of the same construct to drive theoretical reflection, *multilevel/expansion* to investigate different but related constructs in hierarchical multilevel systems, and *evaluation/intervention* to explore different but related constructs in relation to effectiveness.

Two points to note about the classification of MMR according to purpose are the prevalence of certain purposes in empirical studies and the relationship between purposes and certain design features. First, not all rationales for mixing methods are equally present in the literature. Creamer (2018: 30) notes that designs with the purpose of initiation are "underutilized but [...] have] tremendous potential". In addition, methodological reviews indicate that in the social sciences enhancement/complementarity is the most common MMR purpose, although there are discipline-specific differences (Creamer 2018: 32). Second, different purposes are frequently associated with design features: for example, if the rational is (instrument) development or enhancement/complementarity, researchers are likely to use sequential designs, whereas triangulation/confirmatory purpose may involve concurrent data collection, even in a single session. Similarly, priority maybe given to qualitative methods for an evaluation/intervention purpose, whereas triangulation/confirmatory can typically involve a qualitative follow-up for mainly quantitative data (Creamer 2018: 30–31). This discussion highlights that in MMR, as in all other research, the rationale and questions driving the study should motivate all other design decisions, including timing and priority.

Finally, another aspect that needs to be discussed are quality criteria in MMR. Following the emergence of MMR as a distinct third paradigm, the question of assessing quality in MMR studies came to the forefront in the early 2000s. O'Cathain (2010) identifies three approaches to quality assessment in MMR: a generic research approach, which assumes that all research can be assessed using one set of tools, an individual components approach which proposes that each qualitative or quantitative component be assessed using the criteria associated with the relevant research tradition, and a mixed methods approach, which endeavours to assess the quality of MMR as a whole rather than split into individual strands. A first attempt for such an MMR approach to quality was proposed by Teddlie and Tashakkori (2003: 12) who called for "a bilingual nomenclature". Central to their notion of quality in MMR is the term *inference*, which refers

> to an outcome of a study, whether it is derived inductively or deductively. We used the term inference quality to refer to a process that encompasses both internal validity and credibility. We then defined two aspects of inference quality: design quality and interpretive rigor (Teddlie & Tashakkori 2003: 38).

In later work, the authors also introduce the term *meta-inference*. These are "conclusions generated by integrating inferences obtained from the QUAL and QUAN strands" (Teddlie & Tashakkori 2009: 338). Integration here refers to comparing and contrasting, infusing, linking or modifying the two sets of inferences during analysis (Teddlie & Tashakkori 2009: 300), a process that is unique to MMR and can lead to stronger inferences (see also Creamer 2018).

Other scholars who have built on Teddlie and Tashakkori's (2003) initial suggestion include Onwuegbuzie and Johnson (2006), who propose the term legitimation and argue that rather than viewing legitimation as an outcome it should be regarded as an ongoing process occurring at each stage of MMR (see also Onwuegbuzie, Johnson & Collins 2011). These researchers also emphasize the particular challenge of quality in MMR and propose using the relevant

quantitative and qualitative quality criteria in addressing legitimation in the quantitative and qualitative components of MMR and then additionally addressing the legitimation of mixing these methods; especially with regard to meta-inferences based on inferences made from each part of the study. This means that in practice researchers engaging in MMR need a sound knowledge base in assessing the quality of both quantitative and qualitative research to address the legitimation in either strand before carefully investigating the quality of mixing these strands at all steps of the research design. A useful model to aid researchers with the reporting of quality evidence in practice are the "Good Reporting of A Mixed Methods Study (GRAMMS)" guidelines by O'Cathain, Murphy and Nicholl (2008). It presents six guidelines in relation to different steps in the research process and was also used for the present study. [92]

Mixed methods research is by now an established third research tradition with its own characteristics in relation to paradigmatic foundations, research purposes and designs, and quality criteria, which were followed as closely as possible in the present study. In addition to MMR, school-based research forms a second methodological backdrop to this study: since data collection was carried out in schools, practical constraints and issues common in such research (e.g. Harrell et al. 2000; Lonergan & Cumming 2017; Mackey & Gass 2005; Rossiter 2001; Testa & Coleman 2006) also apply to the present study. In navigating the realities of school life during data collection, care was taken to enhance comparability between different data collection sessions, while also trying not to place additional burdens on teachers and students. These efforts are described in greater detail in the descriptions of the two strands in sections 5.3 and 5.4, but first section 5.2.2 presents the MMR design of the present study.

5.2.2 The research design of the present study

The research design of the empirical study on extramural English and its relation to vocabulary knowledge integrates both quantitative and qualitative methods in a cross-sectional MMR design. The rationale for using MMR primarily is to enhance and complement the quantitative findings with qualitative insights, but at the same time the qualitative strand is used to check, clarify and support the results of the quantitative strand. In Creamer's (2018) typology of purposes for conducting MMR the study can thus be placed in the enhancement/ complementarity and triangulation/confirmatory categories In addition, the results of the quantitative strand influenced the design of the interview guide for qualitative data collection; hence, instrument development is also a part of the MMR design, but not its major purpose.

The aims of the project and the research questions presented in section 5.1 reveal that priority is given to the quantitative component of the study and that it uses a sequential design. Using Morse's (1991, 2003) notation conventions the present study thus uses a sequential QUAN → qual design. Creswell et al. (2003) have called this a "sequential explanatory design", the purpose of which "is typically to use qualitative results to assist in explaining and interpreting the findings of a primarily quantitative study" (Creswell et al. 2003: 227, see also Creswell and Plano Clark

---

[92] For further suggestions on evaluating the quality of MMR studies, see for instance the extensive framework proposed by O'Cathain (2010) or the Mixed Method Evaluation Rubric (MMER) by Creamer (2018).

2007, 2011). While this is true to a certain extent in the case of the present study, the qualitative strand is not only used to explain the quantitative results, but to enrich the quantitative description of engagement with EE by adding qualitative information on aspects not addressed in the quantitative component. In line with the enhancement/complementarity purpose described above, the interview data are thus not only used to clarify and support the quantitative findings but to add a new perspective to the study by giving room to the students' emic views on extramural English and related aspects.

Figure 5.1 presents a visualization of the MMR design used in the main study following the guidelines given in Ivankova, Creswell and Stick (2006) with slight modifications. The left side visualizes the research steps undertaken for both data collection and analysis, whereas the right-hand column provides information on the products or outcomes of each stage. The arrows indicate the procedural flow and connections between the different stages of the study including an iterative loop for data analysis before the final integration of quantitative and qualitative results.

In this sequential design the quantitative stage aims at collecting information concerning participants' use of EE, background variables and their vocabulary knowledge. The instruments used in the quantitative strand are a questionnaire (section 5.3.3.1), a structured language diary (section 5.3.3.2) and two vocabulary size tests (section 5.3.3.3). The products of this stage of data collection are mainly numeric data and some text data from open-response items in the questionnaire. These quantitative data are then subjected to a first round of analysis using mainly descriptive statistical procedures to allow first results to feed into the development of the interview guide (see section 5.4.2) for the qualitative strand. In the second qualitative data collection phase data on the teenagers' perceptions of EE and its potential for language learning are collected with the help of focus-group interviews. The products of this stage thus consist in audio data, which are then converted into text data through transcription, and the interview protocols.

In addition to instrument development, the quantitative data are also used to guide the sampling for the qualitative strand. In Teddlie and Yu's (2007) terminology the study uses a sequential sampling strategy in which "[t]ypically, the methodology and results from the first strand inform the methodology employed in the second strand" (Teddlie & Yu 2007: 90). The present study is a typical example of such designs because the larger sample used in the quantitative strand is then used as the sampling frame for the smaller qualitative sample (see sections 5.3.1 and 5.4.1).

Following the collection and transcription of the qualitative data, they are analysed using qualitative content analysis (see section 5.4.6), while concurrently the quantitative data are subjected to further statistical analyses such as regression modelling. At this stage of analysis, results from one strand influenced the analysis in the other, since, for instance, the insights provided by students in the interviews indicated new connections, which could also be explored in the quantitative data.

QUANTITATIVE
data collection
(*N* = 225)

Extramural English
Questionnaire survey
(EEQ)

Extramural English Online
Language Diary (EEOLD)

Receptive and productive
vocabulary tests

QUANTITATIVE
data analysis

QUAN-qual connection

Sample
selection

Instrument
development

Qualitative
data collection
(*n* = 30)

Focus group interviews

Qualitative
data analysis

Integration of results

Product

- Numeric data
  (frequency of EE, background data)
- Text data from open item formats

- Numeric data (EE mean time)

- Receptive vocabulary test score
- Productive vocabulary test scores

- Descriptive statistics
- Correlations and group comparisons
- Regression model

- Interview guide

- Audio data & text data (transcriptions)
- Interview protocol & reflection

- Codes and themes (content analysis)
- Thematic comparisons across groups and participants

- Comparison of quantitative and qualitative inferences
- Meta-inferences
- Relation to previous studies

Figure 5.1: Visualization of the mixed methods design used in the present study

The results of the quantitative and qualitative analysis are at first described separately for each strand (see Chapters 6 and 7) before a final stage of integration. This procedure, which allocated room to data integration from the start, was regarded as helpful in the systematic building of meta-inferences from the inferences drawn in the two strands of the study. The meta-inferences based on the integrated MMR results are presented together with a discussion in relation to previous research in Chapter 8, which thus also integrates the present studies' findings into the wider context.

To sum up, mixing of quantitative and qualitative methods can occur at five stages of a study (Creamer 2018: 12): planning and design, data collection, sampling, analysis and drawing inferences. In the present study mixing is found in the form of the overarching MMR question and the research questions guiding each strand in the planning and design stage. In combination with the aims of the study the research questions also provide a clear rationale as to why an MMR approach is most appropriate for this project. Mixing clearly also occurs at the sampling stage with the interrelated sequential sampling procedure and at the stage of data collection through instrument development. During the analysis the qualitative results sometimes fed back into the analysis of the quantitative data and qualitative data were converted to quantitative data through counting, albeit for descriptive purposes only. Mixing during analysis is perhaps less prominent because the results are reported separately before they are systematically drawn together in the integration phase; however, this final stage clearly entails mixing in the form of combining results and inferences. Following Creamer (2018) the MMR design of the present study can therefore be regarded as fully integrated.

Concerning the assessment of study quality, O'Cathain's (2010) criteria were considered as good practice guidelines in the planning, implementation and reporting phases of the study, but more specific suggestions for sequential QUAN→QUAL designs were also taken into account. Ivankova (2013: 41–42) proposes three strategies to ensure the trustworthiness of meta-inferences: "applying a systematic process for selecting participants for qualitative follow-up, elaborating on unexpected quantitative results, and observing interaction between qualitative and quantitative study strands". All three of these strategies were implemented in the present study through systematic qualitative sampling, gathering focus groups participants' evaluations of unexpected quantitative findings, and keeping an open mind towards mutual influences between the two strands. In addition, the description of the MMR study in this chapter follows suggestions of the GRAMMS model (O'Cathain, Murphy & Nicholl 2008) and aims at being as transparent as possible concerning instrument selection and development as well as data collection and analysis, so that readers are able to judge the quality, suitability and rigour of the study design for themselves.

5.2.3 Research steps including piloting

This section summarizes the overall research procedure followed in the present study including all steps taken to pilot the data collection formats. Due to the sequential timing of the QUAN and qual strand in the present study and the fact that the development of the qualitative instrument was influenced by preliminary quantitative results, both strands were piloted separately. Before the full pilot study of the quantitative data collection phase, several pre-piloting sessions were carried out to determine the target age group and to inform quantitative instrument selection and development.

These pre-piloting sessions took place in the school where I work as a teacher of English. In contrast, the full pilot study was conducted at a different academic secondary school, which met all criteria for inclusion in the main study (see section 5.3.1) and for which permission was kindly granted by the educational board of Vienna (03.06.2016). Participation in all pre-piloting activities and the full pilot study was absolutely voluntary and took place during school lessons.[93] Table 5.1 provides an overview of all steps taken, the pre-piloting and piloting phases are described in more detail below.

Year	Date	Data collection	Purpose
2015	12.11.2015	pre-pilot with retrospective questionnaire with 9 students in 12th grade	determine age group
2016	20.01.2016	pre-pilot with retrospective questionnaire with 17 students in 12th grade	determine age group
	27.04.2016	Pre-pilot for Lex30 with 23 students in 8th grade	determine duration and feasibility of Lex30
	12.05.2016	Pre-pilot for EEQ with 20 students in 10th grade	determine duration and feasibility of EEQ collect feedback on wording of items
	13.06.2016	Pre-pilot for VST with 24 students in 10th grade	determine duration and feasibility of VST
	17.06.2016	Pre-pilot for V_YesNo with 25 students in 9th grade Pre-pilot for revised EEQ with 22 students in 9th grade	determine duration and feasibility of V_YesNo determine duration and feasibility of EEQ collect feedback on wording of items
	14.-21.06.2016	Full pilot study of the quantitative strand with 21 students in 10th grade	determine order of instruments and time needed for administration determine feasibility of EEOLD and get feedback get further feedback on EEQ
2017	23.5.2017	Qualitative pilot of focus group interviews with one group of 5 students in 10th grade	collect feedback on the questions in the interview guide determine duration of the interview and feasibility of the procedure

Table 5.1: Overview of all steps taken to pilot the instruments used in the quantitative and qualitative strands of the main study

---

[93] In order not to disturb regular teaching lessons, all pre-piloting sessions were carried out in cover lessons on days when the regular teachers were absent.

The first empirical step was to determine the best age group to target in the main study. Therefore, pre-piloting sessions were conducted with two groups of 12th-grade students. They were given a short questionnaire with questions on their engagement with EE and retrospective questions on when they had started doing EE activities. Results of the pre-piloting session showed that for these 26 12th-grade students, music was generally the first EE activity. Reading and watching films and series were reported as the last activities that these students began doing with mean starting ages between 13 and 14 years. Hence, the full range of EE activities that the students engaged in in grade 12 was already present at the age of 15 (grade 10) and therefore, it was decided to focus on this age group in the main study. The pre-piloting sessions with 12th-grade students further showed that response options such as "always" and "sometimes" were not feasible for use in a questionnaire and more concrete time indications were preferable. In addition, informal discussions indicated that engagement with EE was highly variable and individualized, which is why I decided to develop a very detailed EE section for the questionnaire.

After a long development phase and extensive discussions with colleagues, first versions of the Extramural English Questionnaire (EEQ) were tried out twice in pre-piloting sessions before the full pilot study. The main purpose of these sessions was to find out whether the questions were comprehensible to students in 9th and 10th grade and whether any of the wordings or concepts used were unclear or could be misunderstood. Therefore, these students were invited to mark questions that were unclear to them with highlighters and/or write comments in the margins. In addition, I also clarified specific questions about concepts such as "leisure time" which teenagers might understand differently from adults. Due to students' comments, changes were then made to the format and wording of several questionnaire items, but the construct (see section 5.3.3.1) remained unaffected. In addition to the EEQ, three vocabulary tests, Lex30, V_YesNo and the VST, were also pre-piloted for inclusion in the main study (see section 5.3.3.3). The purpose mainly was to establish time frames and since students did not grant permission to use their results, they are not reported here.

A full quantitative pilot study with 21 10th-grade students was conducted at a Viennese academic secondary school in the last three weeks of the school year 2015/16. This pilot fulfilled several purposes: first, the complete quantitative data collection procedure could be evaluated; second, feedback on the order of instruments could be gathered and third, the Extramural English Online Language Diary (EEOLD, see section 5.3.3.2) was trialled for the first time. Data were collected in three sessions: in session 1, the students filled in the EEQ and received instructions for the EEOLD; in session 2, they took the VST and gave feedback on the EEOLD, and in session 3, Lex30 and V_YesNo were administered.[94] Unfortunately, in the last session of the pilot study, almost half the class was absent, which was probably connected to the time of data collection in the last weeks of school.

---

[94] The instruments used in the full quantitative pilot study are not included in the Appendix A as only a few and minor changes were made as a result of this final pilot.

One aspect of particular interest in the pilot of the quantitative strand was the time needed to complete the different instruments. In session 1, the students needed 26 to 35 minutes to complete the EEQ, which allowed time to give detailed instructions regarding the EEOLD. Among the vocabulary tests, the VST took most time with 32 to 35 minutes for the majority of students, whereas for V_YesNo only five to six minutes were needed and eight to ten minutes for the additional translation items (see section 5.3.3.3). For the timed administration of Lex30 only 13 minutes were needed instead of the anticipated 15 as all students finished slightly earlier. One insight from the full pilot study therefore was that the administration of both the VST and Lex30 within one 50-minute school lesson was not feasible, given that time would also be needed for the instructions.

Furthermore, the EEOLD was trialled for the first time; students reported taking about three to five minutes to fill the diary in once. They had some comments on the items and their format, especially with regard to the online administration and filling it in on smartphones, but no major changes were made. One issue that surfaced with regard to the EEOLD was that completion rates over the span of one week were relatively low. All students filled in the instrument at least once, but several did not complete a diary entry for every day of the week. This problem was probably due to the fact that for reasons of anonymity I could not regularly remind them to do the language diary (see section 5.3.4.2), but it could have been aggravated by the timing of the pilot study at the end of the school year.

The full pilot study also presented one last opportunity to collect feedback on the EEQ and to clarify students' understanding of concepts, but in general, there were no problems with understanding the items included. The experience gathered during the full quantitative pilot study as well as previous pilot sessions helped to phrase the instructions for the data collection sessions in such a way that frequent questions could be clarified pre-emptively and some problems, such as students turning back to previous test items, could be avoided. In addition, they also served as an opportunity to establish, practise and improve data entry and scoring procedures and thus were a major step towards developing the codebooks and data entry templates used in the main study.

At a much later stage, after the completion of the quantitative strand and the development of the interview guide for the follow-up focus group interviews, the interview guide as well as the general interview procedure were piloted with one group of five 10th-grade students in the school I teach at in May 2017. The participants volunteered to take part in this pilot interview outside of class time and were told that their data would not be used in the research project except for adaptations and improvements to the interview guide. After a short introduction, the procedure of the pilot interview followed the plan for the focus groups as closely as possible to be able to assess the time needed for an unhurried discussion of the interview questions. In addition, the participants in the pilot received a set of extra questions about their experience of the focus group afterwards. These regarded any general comments they had, anything they disliked or found uncomfortable, whether an aspect of the topic was missing and whether any

of the questions were not readily understandable. During this follow-up discussion, the participants were presented again with the main questions and the visual prompts and were asked to comment on their comprehensibility.

The main part of the pilot interview (excluding the additional questions) lasted 51 minutes, which showed that it was feasible to conduct the focus groups during students' lunch breaks or free periods in the main data collection phase. The participants indicated that overall the interview was interesting and that the questions made sense to them. The last question on the relationship between in- and out-of-school English was not clear to them; hence, several formulations were discussed to make it more understandable and the interview guide was changed accordingly. The use of visuals was evaluated very positively and since there were no comprehension problems these were left unchanged. In addition to the constructive feedback from participants, the pilot interview also helped to practise and develop my moderating skills and revealed several questions students had about the interview procedure (such as the language of the interview), which helped to clarify and improve the introduction for participants in the main data collection phase.

Year	Date	Data collection	Strand
2016	23.11.2016	1st data collection session in class SA01	QUANTITATIVE
	02.12.2016	1st data collection session in class SA02	
	07.12.2016	1st data collection session in class SG01	
	12.12.2016	1st data collection session in class SG02	
	14.12.2016	2nd data collection session in class SG02	
	15.12.2016	2nd data collection session in class SG01 1st data collection session in classes SC01 and SC02	
	16.12.2016	2nd data collection session in classes SA01 and SA02	
	22.12.2016	2nd data collection session in classes SC01 and SC02	
2017	10.01.2017	1st data collection session in class SD01	
	17.01.2017	1st data collection session in class SB01	
	24.01.2017	1st data collection session in classes SF01 and SF02	
	25.01.2017	2nd data collection session in class SD01	
	30.01.2017	1st data collection session in classes SE01 and SE02	
	31.01.2017	2nd data collection session in class SE02 2nd data collection session in classes SF01 and SF02 2nd data collection session in class SB01	
	01.02.2017	2nd data collection session in class SE01	
	13.06.2017	Focus group interview with participants from class SA02	qualitative
	16.06.2017	Focus group interview with participants from class SF01	
	19.06.2017	Focus group interview with participants from class SE01	
	21.06.2017	Focus group interview with participants from class SD01	
	26.06.2017	Focus group interview with participants from class SG02	
	27.06.2017	Focus group interview with participants from class SC01	

Table 5.2: Overview of all data collection sessions in the quantitative and qualitative strands of the main study

Simultaneous to the extensive efforts to pilot all instrument and procedures, preparations for the main study began in late spring 2016. The first steps involved contacting teachers in schools fitting the sampling criteria for the quantitative strand (see section 5.3.1), but the search for participants continued at the beginning of the new school year in September 2016. At the same time, approval for the main study was sought from the educational board of the city of Vienna and the ethics committee of the University of Vienna. Having obtained a positive evaluation from the ethics committee (18.07.2016) and the permission to collect data in seven schools from the educational board (20.10.2016), the main study then took place in the school year 2016/17 from the end of November onwards (see Table 5.2 for an overview of all data collection sessions).

In the main study the procedure for quantitative data collection in each class involved three different parts: in a first step, the research project, which was referred to as "Englisch in der Freizeit" [*English in your spare time*] for students, was introduced at the beginning or end of a lesson one or two weeks before the first data collection session. The introduction took between five and ten minutes including time for questions and students were provided with information sheets for themselves and their parents, which also included a consent form. In the first data collection session following this introduction the EEQ was administered and subsequently the instructions for the EEOLD were given to students. In the second quantitative data collection session the two selected vocabulary measures, Lex30 and V_YesNo (see section 5.3.3.3), were completed by the students.

Quantitative data collection lasted from the end of November 2016 to the beginning of February 2017. Subsequently, the quantitative data were prepared for computational analysis (see section 5.3.5) and analysed descriptively to inform the development of the interview guide used in the qualitative strand. All focus groups in the qualitative strand of the main study took place between mid and end of June 2017 in the last weeks of the school year (see Table 5.2). Data collection for the main study was therefore completed in June 2017 and was followed by a second phase of data entry and transcription and an extended period of qualitative and further quantitative data analysis (see sections 5.3.6 and 5.4.6).

Having described the methodological foundations of the study, its overall design and all research steps taken during the pilot phase and the main study, sections 5.3 and 5.4 will present the quantitative and qualitative strand in greater detail.

5.3 The quantitative strand

This section provides more detailed insights into the quantitative strand of the MMR study including the sampling strategy used (section 5.3.1), the final sample of participants (section 5.3.2), instrument selection and development (section 5.3.3), data collection procedures (section 5.3.4) and data preparation and analysis (sections 5.3.5 and 5.3.6).

The quantitative phase is guided by the following four research questions, which were first presented in section 5.1:

> **RQ 1:** *What kinds of extramural contacts do Viennese upper secondary school students report having with English?*
>
> **RQ 2:** *How much time do Viennese upper secondary school students report spending in contact with extramural English?*
>
> **RQ 3:** *What is the relationship between extramural English and the receptive English vocabulary size of Viennese upper secondary school students?*
>
> **RQ 4:** *What is the relationship between extramural English and the productive English vocabulary size of Viennese upper secondary school students?*

Corresponding hypotheses for data analysis were formulated if appropriate: since RQ 1 and RQ 2 are of a descriptive nature, no inferences are made and therefore the statistical procedures do not warrant hypotheses. Concerning RQs 3 and 4 and the relation between EE and vocabulary size it was hypothesized that the frequency of EE activities has a positive impact on both receptive and productive vocabulary size, but that other factors also influence vocabulary size and potentially mediate this relationship. These relations will be explored using correlations and regression models in Chapter 6.

5.3.1 Sampling strategy and criteria

As described in section 5.2.2, this study uses a sequential mixed methods design with the qualitative sample nested in the larger quantitative sample. Therefore, the quantitative sampling procedure provides the basis for the whole study and several sampling criteria were taken into consideration when deciding on the target group for this study: first, characteristics of the Austrian setting such as the school system and curricular goals for the subject English were taken into account. Second, comparability with other EE studies, such as Sundqvist's (2009a) seminal work in Sweden, was considered and third, practical considerations of availability and access also had to be included in these decisions.

Early on a decision was taken to focus on the school type of academic secondary schools (AHS), also referred to as *Gymnasium* in German. An introduction to the Austrian school system as well as more detailed information on academic secondary schools has been provided in Chapter 4, but it is worth highlighting another characteristic of this school type here, as it influenced sampling decisions. Within their focus on general education, academic secondary schools have the possibility to provide more specific content foci, such as science or the humanities from grade 7 onwards. To include students with a wide range of interests in the study, it was therefore decided to concentrate on larger schools which offer more than one specialization, but at the same time highly specialized forms of academic secondary schools with autonomous curricula were excluded to maintain comparability within the sample.

Furthermore, I decided to limit the population to Viennese students. Although initially a comparison between Vienna and another province was considered, such an extension of the project was practically infeasible. This choice could be regarded as a limitation of the study because it can be argued that it limits its generalizability: since Vienna is both the capital and the biggest city in Austria, a range of English activities are on offer, which means that potentially teenagers attending schools in Vienna are exposed to more English than others. However, since

previous studies (see section 3.3.1) have shown that most exposure occurs via online resources rather than offers such as English-language cinemas or theatres, the physical location might not be as important in research on EE as in other types of research. In addition, the results of the nationwide educational standards test for English in grade 8 and the standardized school leaving exam in grade 12 show that Viennese students in academic secondary schools do not differ markedly from those in the other Austrian provinces with regard to their English proficiency, which indicates that the results based on a Viennese sample can be regarded as an approximation for Austria as a whole.[95]

The final step in establishing the target population for the present study was the pre-pilot study discussed in section 5.2.3, which helped to determine the best age group to target in the study. In line with the results, students in grade 10 were chosen with the added advantage that other studies such as Sundqvist (2009a), Sylvén (2004/2010), Olsson (2012) or Versporr, de Bot and van Rein (2011) used similar age groups in other countries.

Having established that the target population for this research project would be 10th-grade students in Viennese academic secondary schools, the quantitative sample was selected by employing quantitative sampling strategies as faithfully as possible. However, fully random sampling was not feasible for reasons of access and the necessary inclusion of complete classes. Therefore, cluster sampling was chosen as the most viable option, like in many school-based studies.[96] Despite these limitations, the sample should be as representative as possible of the population in terms of gender, age, language background, socioeconomic status (SES), and past English teaching and learning. Moreover, it needed to be large enough for the application of inferential statistical procedures to allow conclusions about the population as a whole.

To achieve these aims, a sampling procedure involving several steps was used. In a first step, a list of all Viennese academic secondary schools provided by the Ministry of Education (Bundesministerium für Bildung, Wissenschaft und Forschung n.d.-a) was consulted and the following exclusion criteria were applied to it: first, the funding body had to be the state, private schools were excluded as they often attract a narrower range of students in terms of linguistic, cultural and socioeconomic diversity. Second, the schools had to offer both the lower and upper levels of secondary education (see section 4.2.1). Some Austrian academic secondary schools only offer upper secondary education from grade 9 onwards (*Oberstufenrealgymnasium*), but

---

[95] In the educational standards test in 2013 Viennese students attendings grade 8 of academic secondary schools performed very similar to their counterparts in the eight other provinces and attained a mean score very close to the nationwide mean result (Schreiner & Breit 2014: 87). In the standardized A-level exam for the school year 2016/17 Viennese 12th-grade students attained slightly better, but still comparable grades than candidates in other provinces. In Vienna, 27.3% of the A-level candidates received the highest grade *Sehr gut* compared with the national average of 22.7% and 30.4% achieved the second highest grade *Gut* compared to 29.5% in Austria. 6.3% failed the exam, whereas nationwide 7.4% did not pass (Bundesministerium für Bildung, Wissenschaft und Forschung 2019).

[96] Concerning the role of probabilistic sampling in quantitative research, Döring and Bortz (2016: 300) point out that real random samples are rare in the social sciences and that whole research areas work mostly with non-probabilistic samples and still use parametric statistics to analyse their data. The situation is similar in the field of second language acquisition as random samples are practically impossible to obtain, particularly in school-based research. The present study is no exception and uses a non-probabilistic sample drawn with the help of stratified cluster sampling and convenience sampling.

since students in such schools have attended different lower secondary schools and thus have different educational backgrounds these were excluded.[97] Third, small academic secondary schools offering only one type of content focus were excluded, as outlined above. Fourth, schools providing special forms of English teaching in the form of CLIL, international degrees, such as the Internationale Baccalaureate, or additional English classes were excluded to ensure that all participants had experienced comparable amounts of English teaching as specified in the Austrian curriculum.

After application of these exclusion criteria 33 out of 96 Viennese academic secondary schools were left for potential inclusion in the study, with most schools having been excluded for being run by private organizations (n = 27) or providing special forms of English teaching (n =13). In a next step, several options of drawing a stratified sample from these 33 schools were explored. However, the inclusion of SES as a stratifying variable was not feasible because the relevant data are not publicly available; hence, the number of multilingual students was used to stratify the sample. In the *Schulstatistik* (school statistics, Statistik Austria n.d.) this variable is operationalized as the number of students who indicate that they speak a language other than German at home in a survey at the beginning of each school year. On the basis of the most recent data, the 33 remaining schools were split into three categories: schools with a high proportion of German-speaking students (>60%), schools where the proportion of speakers of German and speakers of other languages were roughly equal (40-60% of German-speaking students), and schools were the majority of students commonly spoke languages other than German at home (German-speaking students <40%). Most of the 33 academic secondary schools had a high proportion of German-speaking (n = 22), eight schools fell into the second category with roughly equal proportions, and only three schools had a proportion of students with an L1 other than German of over 60%.

Following this stratification procedure, two schools from each of the three clusters were selected to be contacted about the study. According to the sampling plan, in each of these six schools two 10th-grade classes, preferably with different specializations, would be invited to participate, resulting in a total of 12 school classes. Using a conservative estimate of class size in Austrian 10th-grade classes, this would result in 240 students being asked to take part. Considering common dropout rates due to lack of student or parental consent, withdrawal of consent, and absences, it was estimated that between 150 and 200 students would complete the quantitative strand of the study. In practice, access to schools is difficult to gain; hence, I selected schools where I had at least one contact person I could personally approach about the possibility of joining a research project. I also contacted teachers first because I considered them to be the most important stakeholders for data collection (see Testa & Coleman 2006). If my contact teachers or their colleagues teaching 10th-grade classes showed interest in the study, I met with them at the schools to discuss the details. These face-to-face meetings were valuable

---

[97] It is for instance likely that some students attending *Oberstufenrealgymnasien* would have previously gone to New secondary schools (see section 4.2.1), which would have introduced additional variables into the study design.

opportunities to discuss the practical aspects of data collection and to allay teachers' concerns regarding additional workload, data protection or the process of gaining official permission. In total, English teachers in eight schools were contacted and in seven of these at least one teacher agreed to allow time for data collection in their lessons. In a next step, I gained support heads of schools as well as their written consent in personal meetings or via telephone calls and e-mails.

In addition, consent also had to be gained from the participants' parents or legal guardians as data were collected from minors. In line with the guidelines of the ethics committee of the University of Vienna special care was taken to provide easily understandable information about the study including details concerning anonymity, voluntariness of participation, data protection, data use and the possibility of withdrawing consent after data collection for both the students and parents. The information sheets and consent forms for students and parents are available in German in Appendix A. In addition to providing this information in print, I personally introduced the project to each class prior to data collection (see section 5.3.4).

School Code	Cluster	Class Code	Total number of students in group
SA	1	SA01	16
		SA02	21
SB	1	SB01	12
SC	1	SC01	28
		SC02	27
SD	2	SD01	19
SE	2	SE01	20
		SE02	20
SF	2	SF01	17
		SF02	16
SG	3	SG01	24
		SG02	19
7 schools		**12 classes**	**239 students**

Table 5.3: Overview of school and classes taking part in the study

In total, students from 12 classes or English groups formed the final sample of the study. As shown in Table 5.3, the classes as well as the schools will be referred to using anonymized codes. Table 5.3 also presents some information regarding the seven schools: three each formed part of sampling clusters 1 and 2, while only one school from cluster 3 agreed to take part. All schools are located in different districts of Vienna with the exception of schools SE and SF, which are both situated in one of the larger districts. The sample includes both highly prestigious academic secondary schools, as well as secondary schools which have a relatively high proportion of students from migrant families, particularly school SG. The sample thus includes students from a variety of backgrounds, but for reasons of anonymity more specific information on the schools cannot be provided. Finally, the number of students in Table 5.3 represents the total number of students in a class or group, but of course students could decide individually not to take part in the study. Further information on all participants in the quantitative strand is provided in the next section.

5.3.2 Participants

A total of 224 students from the seven schools described in the previous section participated in the quantitative strand of the study. However, several of these participants completed only parts of the study or had to be excluded for other reasons: from the results of the Extramural English Questionnaire (EEQ) it became clear that some students spoke English as a home language, had spent a semester abroad in an English-speaking country or had recently changed schools and had previously attended a school with English-medium instruction or CLIL lessons. Since these students had experienced English differently or received different kinds of English teaching, it was decided to exclude them from the sample after data collection.

In addition, during the administration of the questionnaire as well as from its results it emerged that three students had only recently arrived in Austria. Two of these students approached me during data collection because the questionnaire was in German, which they did not yet speak well and for one student it became clear from the information provided in the questionnaire that they had spent less than a year in an Austrian school. Since the study aims to explore the extramural English behaviour of more typical Viennese students attending academic secondary schools, these three students were also excluded from the sample.

The application of these exclusion criteria led to a reduction of the total sample size by 23 students (see Table 5.4), leaving data from a final sample of 201 students to be analysed.

Reason for exclusion	
English as a home language	13
stay in English-speaking country longer than 3 months	5
previous EMI education	2
recently moved to Austria (< 2 years of residence)	3
Total number of excluded participants	**23**

Table 5.4: Overview of participants excluded from the study after data collection

To compare the final sample size to that of the target population, data can be obtained from the school statistics available in the STATcube by Statistik Austria (n.d.). According to these, 5,178 students attended grade 10 in an academic secondary school providing both lower and upper secondary education in Vienna in 2017. The 201 students of the final sample used in the quantitative analyses thus constitute 3.9% of the total population.

Typically, these students are 15 or 16 years old and have begun their English education in the first grade of primary school. At the end of grade 9 they should have achieved CEFR level B1 in English, which should then be consolidated and expanded in grade 10 according to the curriculum (Bundesministerium für Bildung, Wissenschaft und Kultur 2004). Based on standard curriculum regulations, these students have received approximately 805 English lesson, which each last 50 minutes; thus, the total hours of instruction amount to 671 full hours. More detailed information on the participants in the quantitative strand is provided in section 6.1 based on the descriptive results of the EEQ.

5.3.3 Quantitative instruments

This section includes information on the design or selection process for each instrument and on the changes made in response to the pilot studies. The final versions used in the main study are described and are also provided in Appendix A. Timewise, the Extramural English Questionnaire (EEQ) and the Extramural English Online Language Diary (EEOLD) were constructed in parallel to the selection of the vocabulary measures.

5.3.3.1 The Extramural English Questionnaire

Questionnaire are the most widely used research instruments in SLA after language proficiency tests (Dörnyei 2010: xiii). They present a number of advantages in terms of versatility and time effectiveness but have also been severely criticized because they often are poorly designed "ad hoc instruments" (Dörnyei 2010: 1), which lack methodological background and are not grounded in a well-defined construct. Further concerns regarding the use of questionnaires regard the impossibility of remedying missing data and mistakes, problems of respondent understanding or reliability, several bias effects and the relative superficiality of the data. For these reasons, Mummendey and Grau (2014) regard questionnaires as subjective research instruments. They argue that filling in a questionnaire involves several sociocognitive processes such as interpreting a question, retrieving relevant examples from memory, forming a judgment about the perceived question, and positioning this judgement on a given scale. As a consequence, the authors consider questionnaires to be a form of social interaction between the researcher and the participant as respondents interpret and react to linguistic expressions of the researcher much in the same way as in a verbal interview, albeit in written form. Questionnaires thus are sites of co-construction of meaning and participants' responses are a product of their knowledge and experience, of several situational factors and the research instrument itself (Mummendey & Grau 2014). Researchers need to bear the context-specificity of the data given by the instrument, the data collection procedures and participant characteristics in mind when interpreting results.

Clearly, the design and use of questionnaires is not unproblematic, but as Dörnyei (2010: 11) points out "careful and creative construction can result in an instrument that motivates people to give relatively truthful and thoughtful answers, which can then be processed in a scientifically sound manner". For the present study the construction of a new questionnaire was deemed necessary because of its exploratory nature in a relatively new research context (see section 4.4). Although the design process did not start from scratch but drew on examples used in previous research, care was taken to follow advice in the literature (Kirchhoff et al. 2001; Mummendey & Grau 2014), and particularly in Dörnyei (2010), with regard to issues such as the wording of items, response options, item sequence and the construction of scales to design a sound and useful instrument.
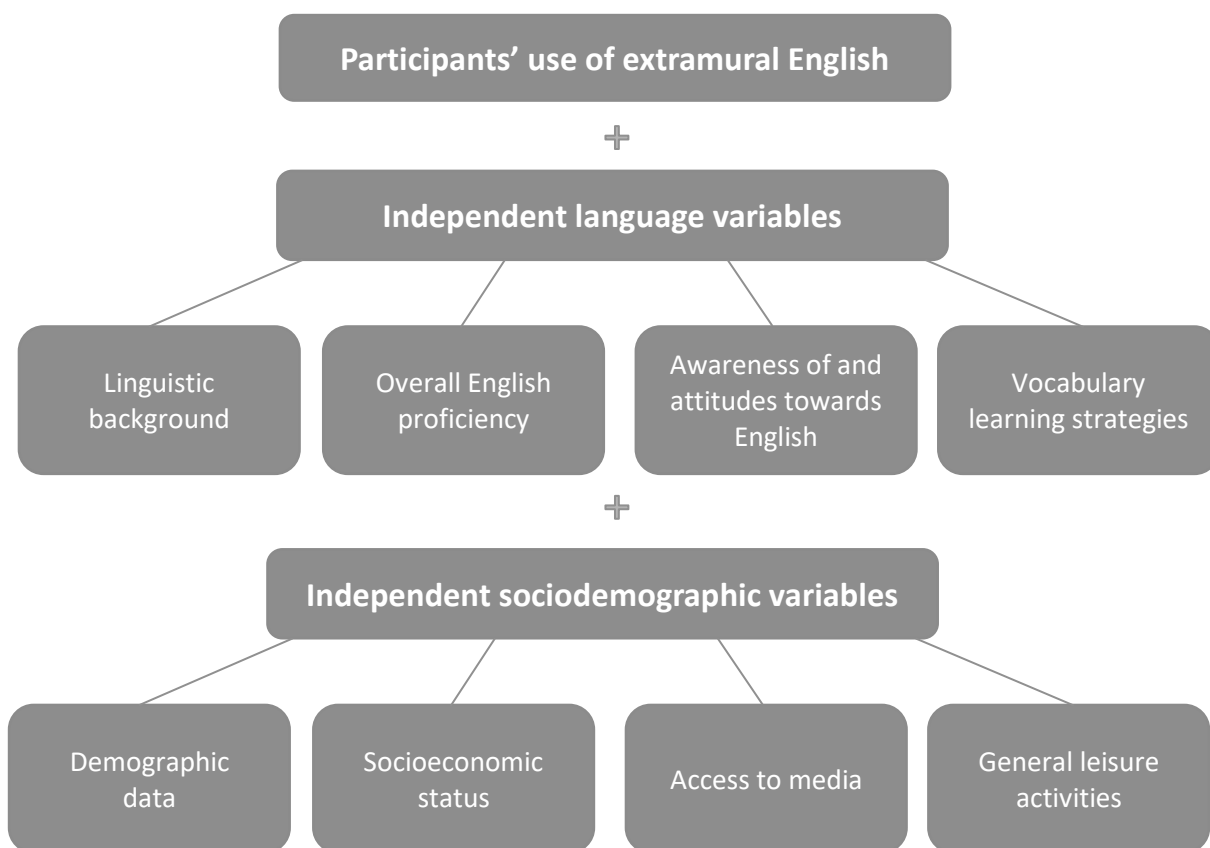
Figure 5.2: The construct of the Extramural English Questionnaire

The purpose of the Extramural English Questionnaire (EEQ) is to elicit information about students' EE practices as well as their linguistic and sociodemographic background. The construct of the EEQ presented in Figure 5.2 thus consists of three main aspects: the participants' use of EE, independent language variables and independent sociodemographic variables. Language variables include students' personal language background, their English proficiency, their awareness of and attitudes towards English, and their use of vocabulary learning strategies. Sociodemographic variables refer to demographic data such as their gender and age, the socioeconomic status of their parents, access to media at home and their general leisure activities and media use within these. Further aspects such as motivation, learning style or linguistic aptitude were considered for inclusion in the construct, but both the length and the content of the EEQ had to stay manageable for participants. The final version of the questionnaire (see Appendix A) consists of seven thematic sections beginning with simple questions on the availability of media and general leisure time activities (1), followed by English leisure activities (2), attitudes towards English (3), participants' linguistic environment (4), strategies for new words (5), participants' language background (6) and a general background section (7) entitled "You and your family". All items are presented in German as the language of schooling to ensure that students understand them correctly and to reduce the amount of time needed to complete the questionnaire.[98]

---

[98] An English translation of the EEQ is available upon request.

In the following, detailed information on the operationalization of variables is provided for each part of the questionnaire. First, existing questionnaires were analysed in parallel to consulting literature on questionnaire design. Table 5.5 shows the three types of instruments that were particularly informative during the operationalization phase; in the following, their use and the process of operationalizing the construct of the EEQ is described for each of the three parts of the construct shown in Figure 5.2.

EE studies and related research	Surveys on the use of English	Large-scale educational studies
S. Bajor, personal communication, 23.11.2015 [99]	Edwards (2016)	DESI: Wagner, Helmke & Rösner (2009)
Berns, De Bot & Hasebrink (2007)	Leppänen et al. (2011)	PISA 2012: OECD (2014), OECD/Bifie (n.d.)
Sundqvist (2009a)		
Sylvén (2004/2010)		
Hyland (2004)		
Lai, Zhu & Gong (2015)		

Table 5.5: Three types of questionnaire studies informing the design of the EEQ

Questionnaire studies on EE and related topics (see column 1 in Table 5.5) were analysed and compared before designing the items eliciting information on participants' engagement with extramural English. Items eliciting information on respondent behaviour are used less frequently in questionnaires than items tapping into latent constructs (Mummendey & Grau 2014), but since observations of actual behaviour are not feasible in the case of EE, questions on specific activities are the best approximation available. For this reason, participants are presented with a comprehensive list of leisure activities in the EEQ and are asked how frequently they do these in English. The items also specify how students engage with these activities; for instance, whether they watch a film on TV, on DVD, or online. These distinctions were included to allow a more fine-grained analysis of the sources of extramural English in Austria. This resulting list (item set 2a) forms the core part of the EEQ; in addition, participants are also asked to name their favourite English-language books, films, series or webpages to gain an insight what is popular among Austrian teenagers (item set 2b).[100] The next item (2c) elicits information on all languages students regularly use in their spare time. The participants are then asked to consider the proportion of these languages in a visualization task (item 2d).[101] Finally, the students are presented with a number of reasons for using EE and are asked in how far they agree with them (item set 2e).

Variables related to language form the second part of the EEQ's construct (see Figure 5.2). Participants' linguistic background includes language(s) spoken at home and with friends, their contact with English in Austria and abroad, and information on learning English and other

---

[99] Stephanie Bajor is a PhD candidate at Augsburg University conducting a survey on the use of extramural English among German secondary students. She kindly provided me with her already finished instrument, which was of great value due to the comparability of our research contexts.

[100] Item numbers given in the text refer to the final version of the EEQ, which can be found in Appendix A.

[101] The idea of analysing use of different languages in terms of proportions was adapted from Berns, de Bot and Hasebrink (2007) and Bonnet (2004).

foreign languages. These items are mainly based on the questions used in international large-scale surveys (see column 3 in Table 5.5) but were also informed by previous EE research. The variables included are the number of languages spoken at home (item 6a), the starting age of learning English (6c), attendance of English language camps (6e), use of English outside school (6b) and during holidays abroad (6f, 6g), longer stays in English-speaking countries (6h) and other languages learned at school (6d).

Overall English proficiency is also considered as an independent variable in the research design to explore its relationship with extramural English. However, unfortunately in Austria no standardized test results are available for grade 10 and it was not feasible to administer a whole battery of tests to determine overall proficiency in addition to the instruments used. Therefore, two approximations are used to collect information on participants' English proficiency: they were asked to report their last grade in the school subject English (item 6j) and they filled in a self-report scale based on the CEFR (item set 6i).[102] The use of such a scale was inspired by Bonnet (2004), but the scale used, which specifies criteria for each of the four skills, was adapted from the global CEFFR descriptors by myself. Although self-assessment is not the ideal operationalization of students' overall English proficiency, it has been used in previous studies (e.g. Leppänen et al. 2011) and in large-scale projects, such as DIALANG (Alderson 2005) or the Special Eurobarometer on languages (European Commission 2012b). In addition, studies investigating student self-assessment have found positive correlations between self-assessed and test-based evaluations of language proficiency (Oscarson 2014: 719).

In addition to students' linguistic background and their overall English proficiency, their attitudes towards and awareness of English in their everyday surroundings are also explored in the present study. Awareness of English in students' surroundings has not yet been included in EE studies, but such an exploratory analysis could lead to interesting findings with regard to students' perceptions of English. Awareness of English was operationalized as one open question and two closed questions based on items used by Leppänen et al. (2011). In the open question (item 1i) students are asked to consider their everyday life and to name the top three places or situations in which they encounter English most frequently. This item was placed in the first, more general section of the questionnaire before the questions on extramural English so as not to narrow participants' focus to their leisure time only. In two closed items (4a and 4b) students are then asked more specifically how often they see or hear English in specific environments and whether they think they use English more in school lessons or in their leisure time.

For similarly exploratory motives it was decided to also include questions on students' attitudes towards English; or more specifically, on their evaluation of five attitudinal constructs with regard to English and its role in their lives. These are:

a) English is important for the future
b) English is vital for international communication and travelling

---

[102] In addition to their last grade participants were also asked which grade they expected to get at the end of school in which data collection took place (EEQ item 6k), but these data were not used in data analysis.

    c)   English is especially important for young people

    d)   English plays a significant role within Austria.

    e)   English is regarded as better than German.

These constructs as well as their operationalizations were informed by previous studies such as the DESI project (Wagner, Helmke & Rösner 2009) and studies on English in Finland and the Netherlands (Edwards 2016; Leppänen et al. 2011, see columns 2 and 3 of Table 5.5). For each of the constructs four to five statements were formulated to form five multi-item scales (see e.g. Dörnyei 2010). These are presented as one set of items in the EEQ (item set 3) and participants are asked to indicate the extent to which they agree.

The final language variable considered in the EEQ corresponds to the study's focus on vocabulary and taps into students' vocabulary learning strategies as well as their attention to new words in EE input. The latter variable was intended only for descriptive purposes and operationalized as five statements expressing different levels of attention to unknown English words encountered during EE activities (item set 5a). In a second set of questions (item set 5b) the participants are asked to indicate what they do when they encounter a new English word outside school lessons. The options provided are based on the discovery strategies in Schmitt's (1997) taxonomy of VLS. As discussed in section 3.1.3, VLS taxonomies can be considered outdated nowadays, but the newer tendency towards self-regulation (Tseng, Dörnyei & Schmitt 2006) could not be taken into account due to the complex nature of this construct, which would have warranted an instrument of its own. In addition to the set of discovery strategies in item set 5b, students are also asked whether they do anything to check hypotheses they have about unknown words' meanings (item 5c) and whether they do anything specific to memorize such new words (5d) in two open items.

In addition to language variables, the construct of the EEQ (see Figure 5.2) also includes other independent variables relating to participants' sociodemographic background. General demographic information is collected in the final section of the questionnaire including participants' gender (item 7a) and their year and country of birth (7b). Furthermore, this section includes questions on the socioeconomic status of participants and their families. Measuring SES can be difficult as pointed out by Brese and Mirazchiyski (2013: 37): "Socioeconomic status (SES) is by far the most prominent and widely used latent construct for measuring family background. It is also the least well-defined concept". According to Mueller and Parcel (1981), SES can be conceptualized as an individual's or family's position on a social hierarchy based on the dimensions of wealth, power and prestige (see also Brese & Mirazchiyski 2013; Caro & Cortés 2012; Hansson & Gustafsson 2013). In sociology, SES is therefore regarded as "a composite of standing on occupation, education, income and other status dimensions" (Mueller & Parcel 1981: 25) and thus family SES in educational studies is often operationalized as parental education,

occupation and income or financial resources (Brese & Mirazchiyski 2013: 13).[103] What is clear is that SES is an important influencing variable because "[t]here is an empirically established relationship between academic outcomes and SES" (Hansson & Gustafsson 2013: 149). Its immense educational impact even led the American Psychological Association (APA, n.d.) to call for an inclusion of SES in all education research.

The present study attempts to follow this call despite the difficulties involved in operationalizing SES.[104] As mentioned above, one way to operationalize it is to measure the related variables of parental education, occupation and income like in large scale studies such as PISA (OECD 2014). In the PISA student questionnaire home possessions are used as substitute for family income (Brese & Mirazchiyski 2013: 41).[105] Using home resources as an additional indicator for SES has been a common approach since the end of the 1980s (Brese & Mirazchiyski 2013: 24) and has also been adopted in the present study since questions about income can cause discomfort and can probably not be answered by students, whose reports are the only source of information. Studies have however found that adolescents are able to provide relatively reliable information about parental education and occupation (e.g. Lien, Friestad & Klepp 2001). SES is included in the EEQ in form of parental education (item set 7c), parental occupation (items 7d-7g) and family affluence operationalized as home possessions. Home possessions are tapped into in five items: the number of books at home (7h), the number of phones, TV sets and computers available at home (1a), the availability of a range of additional media devices (1b), student-owned possessions (1c) and the availability of an internet connection at home (1d). These items are modelled on the student questionnaire used in the 2012 PISA study in Austria (OECD/Bifie n.d.) with slight adaptations.[106]

In addition to providing information on participants' socioeconomic background, EEQ items 1a to 1d are also used to measure access to different media in students' homes. This is the third independent variable in this category of sociodemographic variables as the availability of different media devices such as TV sets, computers, tablets, DVD players, gaming consoles or e-book readers influences students' access to EE. Previous studies (see sections 2.3 and 3.3) have shown that much EE input is received via popular media and many EE activities are carried out online; hence, internet access is particularly crucial.

---

[103] There is considerable discussion in the literature on how SES should best be conceptualized and in particular how it relates to Bourdieu's concepts of social and cultural capital, which some scholars see as related to SES (e.g. Hansson & Gustafsson 2013), but which others regard as a different model (e.g. Caro & Cortés 2012).

[104] Useful guidance is provided by the extensive discussion of measurement issues in Brese and Mirazchiyski (2013).

[105] Information on income is only gathered in the parent questionnaire in the PISA study, which many countries, including Austria, do not administer.

[106] The question on the employment status of parents (e.g. unemployed, part-time, full-time) was deemed less relevant for the present study and was excluded as it might make some students feel uncomfortable. In addition, the list of home possession was limited to media devices relevant for EE activities and cars were excluded because in Vienna possession of a car is not a reliable marker of SES due to the wide availability of good public transport. Instead, students were asked to indicate whether they have their own room, which, due to the costs of housing, is not always the case in Vienna.

The last variable taken into account in relation to students' sociodemographic background relates to participants' general leisure time preferences. Leisure time activities and media use during students' spare time are of crucial interest in this study because evidently a teenager who never plays digital games in their free time will also not do so in English. For this reason, the general leisure time activities presented in item set 1h serve as a sort of control items for the more detailed list of EE activities in item set 2a. The list of students' general leisure time activities is complemented by three further items asking about the devices commonly used to access the internet (item 1e), daily internet usage (1f), and students' favourite websites (1g).

The development of the EEQ greatly benefitted from constructive feedback by colleagues at the department of English, fellow students in PhD seminars and practising English teachers. In addition, feedback on a revised version was kindly given by Pia Sundqvist, whose experience with EE research resulted in several further changes. In addition, feedback was also collected from students in 9th and 10th grade (see Table 5.1 in section 5.2.3), which helped to identify instructions and items that could be misunderstood or were unclear to students in the target age group. In the final step of development, the administration of the EEQ in the full quantitative pilot study was used to finalize all item wordings, as mentioned in section 5.2.3. In addition, participants in the quantitative pilot study reported that they found it difficult to think about their spare time only; thus, as a final change, reminders to only consider activities in English and in their free time were placed throughout the relevant section.

5.3.3.2   *The Extramural English Online Language Diary*

The Extramural English Online Language Diary (EEOLD) complements the data gathered with the EEQ because it provides information about the time spent with EE rather than the frequency of EE activities. Moreover, the data are more specific as they depict students' actual engagement with EE. Participants are asked to fill in the structured language diary every day for one week; thus, ideally, each student completes the same diary seven times. The EEOLD was modelled on language diaries used in previous EE studies by Sundqvist (2009a), Olsson (2012) and on the language diary used in the CLISS project by Liss Kerstin Sylvén (personal communication, 7.10.2014).

In terms of format, the EEOLD is a short questionnaire that collects information on the use of English during students' leisure time on a specific day (see Appendix A). It consists of seven parts: in the first section the students select the day of the week for which they want to fill in the diary and enter a date, the next five sections consist of items about the amount of listening, reading, gaming, writing and speaking done on this day, and in the last section participants are asked whether they also used languages other than German and English and whether they spent more or less time or the same amount of time as usual with English.

In contrast to the EEQ, in which the EE section is organized according to different media, the EEOLD is structured according to language skills. For instance, the second section on listening begins with the question "Have you listened to English during your leisure time on Monday? To

what and for how long?"[107] Participants are then provided with a list of options for which they can choose "no" as an answer or a time estimate ranging between "about 5 minutes" and "about three hours or more" (see Figures 5.3 and 5.4). Each list of EE activities also contains the option "another activity" and if students select a time estimate for this item, they are presented with a text field and asked to describe the activity. In addition to the list of skill-based activities each section also includes a reminder to only think about their spare time and about English activities at the end. The structure of the reading, gaming, writing and speaking sections is the same with the exception that the section on listening also includes one additional question on the use of subtitles.

Like the EEQ, the EEOLD was administered in German to avoid problems of comprehensibility; moreover, it was provided online to allow participants easy access and to avoid the necessity of involving teachers in the data collection process (for further details on data collection see section 5.3.4.2). The online administration had some impact on the structure of the diary as it allowed a conditional presentation of items depending on participants' responses to previous items; for instance, students were only presented with the question whether they had used subtitles if they reported watching a video clip, film or series on that day.
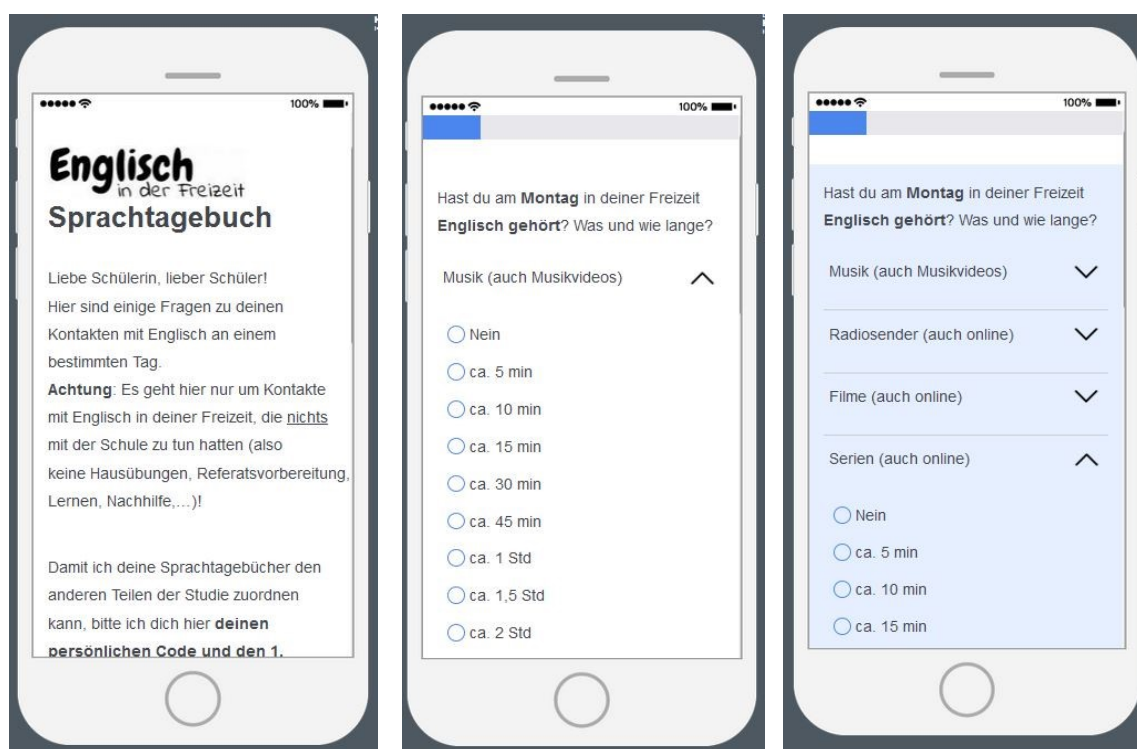
Figure 5.3: Layout of the EEOLD when accessed from a mobile phone

The online platform Qualtrics (Qualtrics LLC 2016) was chosen for the administration of the EEOLD because it automatically adapts the layout of the diary to the device it is accessed from. Figure 5.3 gives an impression of the final version of the EEOLD on a mobile phone; whereas

---

[107] The day in the question is automatically substituted with the day of the week for which students fill in the diary.

Figure 5.4 shows the layout on a computer or tablet. Appendix A includes a list of all items contained in the final version of the EEOLD.

Figure 5.4: Layout of the EEOLD when accessed from a computer or tablet

Similar to the design procedure for the EEQ, feedback on early drafts of the EEOLD was collected from many colleagues, and in addition the EEOLD was tried out in the quantitative pilot study. Participants had few comments on the instrument, but completion rates emerged as a problem in the pilot study because not all students filled in the diary for a full week (see sections 5.2.3 and 5.3.4.2).

5.3.3.3   Vocabulary tests

Researching the relation between extramural English and vocabulary knowledge poses a conceptual problem since the research object 'vocabulary knowledge gained through exposure to EE' is highly individual and cannot be established across all participants. Due to the wide and varied nature of EE input it is impossible to establish specific target items and test participants' knowledge of these. Therefore, the only way to investigate the relationship between engagement with EE and vocabulary knowledge for a larger sample is to use measures that provide estimates of total vocabulary size and analyse how these relate to participants' EE practices.

Previous studies on the link between EE and vocabulary knowledge constituted the starting point for the selection of vocabulary size tests in the present study. Several studies reviewed in section 3.3.1 have explored this link among European teenagers. Two of these studies (Berns, De Bot & Hasebrink 2007; Verspoor, De Bot & Van Rein 2011) used Yes/No tests by Meara and colleagues (Meara & Buxton 1987; Meara 1992), Sundqvist (2009a) used the VLT and PVLT to measure receptive and productive size, Sylvén (2004/2010) employed a combination of measures including the VKS, a MC test, a words in context section and a cloze test, and Peters (2018) used

the VocabLab Test (see section 3.2.3 for descriptions of these size tests). All of these studies thus employed measures of written receptive vocabulary size, which is hardly surprising as this aspect tends to receive most attention with regard to measurement (see section 3.2). In addition, Sundqvist (2009a) and Sylvén (2004/2010) also used tests of productive size or depth.

Because measuring receptive knowledge of spoken word forms using Aural Lex was not a viable option (see footnote 48 in section 3.2.3), I decided to concentrate on written vocabulary size, but to include both receptive and productive size measures. Furthermore, tests producing an overall size estimate were regarded as preferable for statistical reasons. Hence, for receptive vocabulary size the option of using the VLT was not further explored because the level scores are not intended to be added to a sum score, although this is frequently done in research. As a result of this decision, the Vocabulary Size Test (VST), a multiple choice test of the 14,000 most frequent word families by Nation and Beglar (2007), and V_YesNo, the latest in a series of checklist tests by Paul Meara (2015a) testing the 10,000 most frequent lemmas, were selected for piloting. Both measures are not without problems (see section 3.2.3), but as has become clear the perfect vocabulary size test simply does not exist (yet).

For productive vocabulary knowledge the choice of test was more limited from the outset: once the idea of using free production measures was discarded for reasons of comparability and practicality, only Lex30 (Meara & Fitzpatrick 2000) and the PVLT (Laufer & Nation 1999) remained as options. However, similar to its receptive counterpart the VLT, the results of the PVLT should be presented according to levels rather than as one overall size estimate and the test has been criticized for measuring a different form of receptive rather than productive vocabulary (Read 2000). Moreover, early on an idea was formed to analyse the samples of response words elicited by Lex30 in relation to students' vocabulary input at school in addition to the conventional frequency-based scoring method (see section 6.4.7). For this reason, only Lex30 was piloted as a measure of productive vocabulary size.

Hence, in total three vocabulary tests were administered to students in the pre-piloting sessions and the full quantitative pilot study: the VST, V_YesNo and Lex30 (see section 5.2.3). Since the latter two tests are available as online versions from Paul Meara's lognostics website (http://www.lognostics.co.uk/), they first had to be adapted into paper-and-pencil versions (see also section 5.3.4.1). For Lex30 that meant that the 30 cue words given in the appendix of Meara and Fitzpatrick (2000) were inserted into a table with one cue word and four spaces for response words in each line. The layout is presented in Figure 5.5, but the full Lex30 test can be found in Appendix A.

Schreibe neben jedes Wort, die ersten anderen englischen Wörter, die dir dazu einfallen.

1. **attack**			
2. **board**			
3. **close**			
4. **cloth**			

Figure 5.5: Layout of the paper-and-pencil version of Lex30

In the case of V_YesNo, Paul Meara kindly sent the list of words used in the online version of the test (personal communication, 08.04.2016). For scoring reasons, the 100 target items and 100 pseudowords are split into ten groups with each group containing ten target words and ten pseudowords (see section 5.3.5.3). In the paper-and-pencil version these ten groups were aggregated into five blocks and each item was presented together with a "Yes" and a "No" box for participants to tick (see Figure 5.6 and the full version of the test in Appendix A). In addition, 20 target words were chosen and included in form of translation items at the end of the V_YesNo test. These translations items can be used to check whether participants actually know the meaning of words for which they ticked the "Yes" option.

Ich weiß, was das Wort heißt:

Block A			Block B		
acute	☐ Yes	☐ No	bibby	☐ Yes	☐ No
podiast	☐ Yes	☐ No	liverick	☐ Yes	☐ No
malicious	☐ Yes	☐ No	flautism	☐ Yes	☐ No
fair	☐ Yes	☐ No	greenaway	☐ Yes	☐ No
adjoin	☐ Yes	☐ No	appreciate	☐ Yes	☐ No
makeshift	☐ Yes	☐ No	wood	☐ Yes	☐ No
grudgingly	☐ Yes	☐ No	rumour	☐ Yes	☐ No
intimant	☐ Yes	☐ No	allaway	☐ Yes	☐ No
elphick	☐ Yes	☐ No	snape	☐ Yes	☐ No

Figure 5.6: Layout of the paper-and-pencil version of V_YesNo

For the third test, the VST, a new layout was produced based on the version available from Paul Nation's homepage (https://www.victoria.ac.nz/lals/about/staff/paul-nation#vocab-tests). In fact, there are two versions of the VST: an earlier version testing the 14,000 most frequent word families (Nation & Beglar 2007) and two later parallel versions testing the 20,000 most frequent word families (Nation 2012b). In the pilot study the 14K version was used because it has a better sampling rate and because 10th-grade students are unlikely to know words beyond the 14K level.[108] The items were not presented in accordance with frequency levels in my paper-and-

---

[108] However, the item wordings in the 14K version were adapted to those of the newer 20K versions because they use higher frequency words and the response options in the newer version are more equal in length.

pencil version to avoid frustration (Nation 2012b) and two versions were created to prevent students from copying each other's answers. Figure 5.7 presents the beginning of version A used in the full quantitative pilot study.

```
SEE: They saw it.                        PIGTAIL: Does she have a pigtail?
a. close it tightly                      a. a rope of hair made by twisting bits together
b. waited for it                         b. a lot of cloth hanging behind a dress
c. looked at it                          c. a plant with pale pink flowers that hang
d. started it up                            down in short bunches
                                         d. a lover

SOLDIER: He is a soldier.                LATTER: I agree with the latter.
a. person in a business                  a. man from the church
b. person who studies                    b. reason given
c. person who uses metal                 c. last one
d. person in the army                    d. answer

INPUT: We need more input.               EMIR: We saw the emir.
a. information, power, etc. put into     a. bird with long curved tail feathers
   something                             b. woman who cares for other people's
b. workers                                  children in eastern countries
c. artificial filling for a hole in wood c. Middle Eastern chief with power in his land
d. money                                 d. house made from blocks of ice
```

Figure 5.7: Layout of the paper-and-pencil version of the VST (Version A)

In addition, a decision was made to include a short post-test questionnaire at the end of each vocabulary testing session (see V_YesNo in Appendix A) following ideas put forward in Nation (2007: 36–37). Gathering data on how students felt on the day of data collection and how they found the test was considered especially useful since in the full pilot study the vocabulary tests were administered in two separate sessions on different days.

In the full quantitative pilot study in June 2016, 19 out of 21 students were present in the data collection session in which the VST was administered, but unfortunately only 11 students were present for V_YesNo and Lex30. Results show that estimates of receptive vocabulary size based on the VST ranged between 5,300 and 9,800 word families with an average of 7,526 (SD = 1203.16). In contrast, vocabulary size estimates based on V_YesNo ranged between 2,780 and 9,136 lemmas with a mean size of 4,938 words (SD = 1920.26). If the scores of three participants who produced more than 15 false alarms are discarded (see section 5.3.5.3), the mean estimate is 5,119 lemmas (SD = 1901.33). These findings were surprising because the estimates based on the multiple-choice test, in which students have to demonstrate at least partial knowledge of word meaning, are considerably higher than those based on the checklist test, in which students do not have to support their responses with evidence.[109] Further investigation revealed that the number of German cognates in the VST seems to be disproportionately high in relation to the total number of items on the test: myself as well as two other English teachers rated all 140 VST

---

[109] Please note that the difference between the two results is magnified by the two different units of counting used: while the VST uses the larger counting unit of word families (see section 3.1.1), V_YesNo uses lemmas. According to Milton (2009: 12) one can multiply scores based on word families by 1.6 to receive a very rough estimate of the same results in lemmas. This would mean that the mean result of 7,526 word families roughly equate to 12,042 lemmas, which is more than double as much as the mean V_YesNo result of 5,119 lemmas (with FA rates higher than 15 FAs excluded).

items on their cognate status and all three raters agreed that 41 items or 29% can be considered full cognates that students in grade 10 would definitely recognize. In addition, each rater identified several further target words with potential cognate status that might have influenced students' response behaviour in the pilot study. While this clearly is not a thorough study of the factors influencing the VST results in the full quantitative pilot study, this follow-up indicates that English-German cognates seem to be an issue exerting an undue amount of influence in the VST (see section 3.2.2).

Based on the piloting results it thus became clear that using the VST with Austrian learners of English potentially leads to large overestimations of vocabulary knowledge due to the cognate facilitation effect (see section 3.1.3). In addition, I had already received feedback from the first teachers and heads of school I had contacted in preparation for the main study that they were not willing to allow three data collection sessions (see sections 5.3.1 and 5.3.4). Since the VST took up one school session on its own with about 40 minutes testing time, this was the second reason why using the VST in the main study became infeasible. As a result, it was decided not to include the VST in the main study, but to use V_YesNo and Lex30 as the two vocabulary measures.

In the following, these two measures of receptive and productive vocabulary knowledge will be discussed in more detail beginning with V_YesNo. Yes/No tests were first used in L1 research in the 1930s, but it was only in the early 1980s that Anderson and Freebody (1981, 1983) introduced a defining characteristic of this test form: they found that solely relying on learner self-reports produced unreliable results and thus added pseudowords to be able to adjust results for guessing. Meara and Buxton (1987) first transferred this testing method from L1 to L2 research and in the following years Meara and colleagues developed several widely-used checklist tests,[110] but the method has however also been used in other projects such as DIALANG (Alderson 2005) or the LexTALE test (Lemhöfer & Broersma 2012).

In terms of construct, a Yes/No test measures the most basic aspect of word knowledge with the underlying assumption that "if a learner cannot recognize an item as a word in a specific language, it is unlikely that the learner can do anything else with the word" (Harsch & Hartig 2016: 4). In relation to the four constructs describing knowledge of the form-meaning link (see section 3.2.1), Schmitt (2010: 199) states that Yes/No tests "probably should be considered meaning-recall items, even though the meaning does not have to be demonstrated". As pointed out in section 3.2.3, the construct definition also relates to the instructions used in the test: if, as is the case in V_YesNo, test takers are asked to tick words whose meaning they know, the test taps into meaning recall; but if test takers are told to tick the words they recognize, as is the case with LexTALE, the construct more closely resembles form recognition.

The construct of a specific test is thus closely linked with design decisions in relation to format (see section 3.2.1). In general, a Yes/No test merely presents test takers with a list of words and thus "uses the simplest possible format for assessing receptive lexical knowledge" (Beglar &

---

[110] Examples of Yes/No tests developed by Meara and colleagues are the Eurocentres Vocabulary Size Test (EVST, Meara & Jones 1988, 1990), the EFL Vocabulary Test (Meara 1992, 2010), and X-Lex (Meara & Milton 2003).

Nation 2014: 173). As mentioned above, the list of words includes pseudowords to allow adjusting the score for guessing; these pseudowords do not actually carry meaning in the English language, but follow its morphological and phonotactic rules (Beeckmans et al. 2001: 236).[111] Based on this combination of target words and pseudowords in the checklist there are four possible combinations of items and responses shown in the matrix in Figure 5.8.

Figure 5.8: Item-response matrix for Yes/No tests adapted from Beeckmans et al. (2001: 237). The lighter colour indicates correct responses, whereas the darker colour indicates false responses.

As can be seen from Figure 5.8, there are two kinds of correct responses on the test: ticking 'Yes' for a target word and ticking 'No' for a pseudoword. These are conventionally termed a hit and a correct rejection. By analogy, there are also two types of wrong responses: choosing 'No' for a target word and 'Yes' for a pseudoword, which are usually called a miss and a false alarm (FA). It is conventional for scoring to be based on all 'Yes' responses; thus, on the number of hits and false alarms (see below). In this specific format, then, word knowledge "is indexed by the number of hits, usually adjusted by the number of false alarms" (Mochida & Harrington 2006: 79). Milton (2009: 72) calls this checklist format "deceptively simple" because learners are faced with a difficult decision in cases where they are not sure whether they know the meaning of a target word. Much therefore depends on the specific design of the test, including the number of items, the ratio of target and pseudowords, and the instructions given to test takers (Harrington & Carey 2009).

The most important influencing factor in relation to the construct of the test is the exact wording of the instructions. In early research little attention was paid to the influence of instructions on test performance (see Beeckmans et al. 2001 for a critique), but a study by Eyckmans (2004) showed that more specific instructions led to a decrease in false alarms.[112] Interestingly, this was the case although both sets of instructions compared by Eyckmans included a warning about the presence of pseudowords, while earlier tests (e.g. Meara 1992; Meara & Jones 1988) had not alerted test takers to these. By now it seems to be generally accepted that learners should be told

---

[111] In accordance with Beeckmans et al (2001) I prefer the term pseudowords to non-words because since these words follow the rules of English word formation they could potentially exist.

[112] At the same time, more specific instructions did not enhance concurrent validity with a translation test (Eyckmans 2004). Nonetheless, researchers like Schmitt (2010: 200) have argued for the need to use of more specific definitions of "knowing a word" in the instruction of Yes/No tests and several studies (e.g. Fairclough 2011; Harsch & Hartig 2016) have used detailed instructions for Yes/No tests.

about pseudowords and that elaborating on instructions is beneficial, as indicated by the following passage taken from the description of V_YesNo by Meara and Miralpeix (2017: 116–117):

> We usually find it helpful to emphasise to test takers that they should say YES only if they know the meaning of the target word: familiarity with a word form is not enough. We usually tell them that the test is not timed, and there is no advantage in doing the test as quickly as they can. Nevertheless, if they find themselves hesitating over a word, having to think whether they know its meaning or not, then they should answer NO [...] Guessing is best avoided as saying YES to pseudo-words negatively affects the final score. Experience suggests that the test works better if test takers are told that some of the words are not real words, and that they should not answer YES to items they do not know.

Knowing about the presence of pseudowords certainly discourages guessing, which leads to lower FA rates and less overestimation of vocabulary knowledge on the test. Another suggestion to reduce overestimation is put forward by Mochida and Harrington (2006), who told participants that they would be tested on their actual knowledge of some of the items after taking the Yes/No test. The present study used a combination of the recommendations by Eyckmans (2004), Meara and Miralpeix (2017) and Mochida and Harrington (2006). The written instructions for V_YesNo told students to tick 'Yes' if they knew the meaning of the word and 'No' if they didn't. In addition, they were warned about pseudowords and that points would be subtracted if they ticked 'Yes' for a non-existent word form, and that their actual knowledge of the word meanings would be checked for some items after the test. They were also told orally that for this reason it was better to tick 'No' if they were not sure about a given word.

In addition to test instructions, Yes/No tests may also vary in their length, which can affect the scoring procedure. In the EFL vocabulary test (Meara 1992, 2010) each of the levels testing knowledge of 1000 words is represented by 60 target and 40 pseudowords, whereas X-Lex (Meara & Milton 2003) consists of 120 items to measure the most frequent 5,000 words and V_YesNo (Meara 2015a) includes 200 items to test knowledge of the 10,000 most frequent words. Beeckmans et al. (2001: 240) report a recommendation by Meara that a Yes/No test should contain at least 150 items, but ideally more. In addition, the proportion of target and pseudowords also influences the results, particularly if the same scoring formula is compared across tests with a different ratio (Beeckmans et al. 2001; Eyckmans 2004).

Scoring procedures in general are the issue that has been discussed most extensively in relation to Yes/No tests. Conventionally, the number of false alarms, i.e. the number of pseudowords for which participants ticked "Yes", is used to adjust the number of hits to avoid overestimation. As pointed out by Schmitt (2010: 201) there are two ways in which FAs can be used: either test takers who surpass a preset maximum threshold of FAs are excluded as unreliable, or the number of FAs is used to adjust the raw number of hits downwards. The first option is used in a study by

Schmitt, Jiang and Grabe (2011), but most studies opt for the second option using a correction formulae or use a combination of both options, as is the case in the present study.[113]

Name	Formula
h (Number of correct responses)	h
h-f (Hit rate minus false alarm rate)	$h - \text{f}$
Correction for guessing (*cfg*) (Anderson & Freebody 1983)	$\text{P*(h)} = \frac{h - f}{1 - f}$
Δm (Meara 1992b cited in Huibregtse, Admiraal & Meara 2002)	$\Delta m = \frac{(h - f)}{(1 - f)} - \frac{f}{h}$
I_{SDT} (Huibregtse, Admiraal & Meara 2002)	$I_{SDT} = 1 - \frac{4h(1 - f) - 2(h - f)(1 + h - f)}{4h(1 - f) - (h - f)(1 + h - f)}$
S-shaped logistic weighting function (Meara & Miralpeix 2017)	$\text{Vsize} = \Sigma \left(h_s * 100 * (1 - \frac{w(f_s)}{w(h_s)}) \right)$

Notes: h stands for the observed hit rate and f refers to the false alarm rate, $w(h)$ and $w(f)$ refer to the weighted values of the hit and false alarm rate as set out in Meara and Miralpeix (2017: 120).

Table 5.6: Overview of scoring formulae for Yes/No tests

Over the years several scoring formulae ranging from simple counts to elaborate equations have been proposed; Table 5.6 provides an overview of these. As mentioned above, all formulae are based on the number of hits and false alarms with the simplest option being to simply count the number of hits (h). The second option is to subtract the number of false alarms from the number of hits (h-f), which is the recommended formula for X-Lex, for instance. Third, studies (e.g. Meara & Buxton 1987) have also used the correction for guessing formula (*cfg*) suggested by Anderson and Freebody (1983), which is based on the probabilities of blind guessing. In 1992, Meara (1992b cited in Huibregtse, Admiraal & Meara 2002) introduced a new, more sophisticated formula, Delta m (Δm), which is based on Signal Detection Theory.[114] Continuing this line of work, Huibregtse, Admiraal and Meara (2002) proposed a second index based on Signal Detection Theory (I_{SDT}) because issues had been identified with Δm. Finally, the newest Yes/No test by Paul Meara, V_Yes/No (Meara 2015a), uses a completely new scoring formula which is based on an S-shaped logistic weighting function (Meara & Miralpeix 2017). In contrast to previous suggestions, which use the overall FA and hit rate for computation of a score, this new formula requires a different procedure: the 200-item Yes/No test is first split into ten sections and a score using the weighted hit and false alarm rate is calculated for each section before adding the scores of all sections up to obtain the overall estimate of vocabulary size.

---

[113] For instance, a reliability threshold of 10 false alarms out of 20 pseudowords was set for the EFL Yes/No Vocabulary Tests (Meara 1992), although the test uses a correction formula (Eyckmans 2004: 44). For further discussion and an empirical investigation of maximum FA thresholds, see Stubbe (2012).

[114] For a discussion of Signal Detection Theory as well as a comprehensive overview of scoring formulae, see Huibregtse, Admiraal and Meara (2002).

Studies have also compared Yes/No tests to other measures: Mochida and Harrington (2006) compared scores on the VLT to a Yes/No test with the same 90 target words plus 60 pseudowords and found that on average the raw number of hits was the best predictor for performance on the VLT, but all scoring methods were strongly related to the VLT scores (r ranged between .85 and .88, p <.001). Eyckmans et al. (2007) investigated performance on a computerized Yes/No test in comparison to a translation format with the same target words and found lower correlations ranging between r = .663 and r = .741. Harsch and Hartig (2016) compared the predictive power of X-Lex and a contextualized C-test for reading and listening scores in a large-scale study with German students. They found that using separate scores for hit rate and FA rate in their statistical models produced better results than the previously proposed adjustment formulae, but the C-test had higher predictive power overall. In their study hit and FA rate were positively correlated and familiarity of vocabulary, operationalized as higher frequency, as well as higher overall language proficiency lowered the use of guessing.

In an innovative study Pellicer-Sánchez and Schmitt (2012) introduced a timed Yes/No test and compared results based on reaction times to previously proposed correction formulae.[115] 50 L1 and 55 L2 users of English took the Yes/No test, which consisted of 40 target words and 16 pseudowords. In addition, personal interviews were conducted to establish participants' actual knowledge of target words' meanings. Correlational analyses show that for L1 users h-f and the reaction time approach showed the highest correlation with the interview scores, whereas for L2 users h-f produced the highest correlation followed by I_{SDT} and Δm. Overall, Pellicer-Sánchez and Schmitt (2012: 504) conclude that "[t]he effectiveness of these adjusting approaches seems to depend on both the FA rate and the size of participants' overestimation" and thus suggest that ideally one should use different correction formulae for different FA rates. Another important finding of this study is that the Yes/No test score showed considerable overestimation of word knowledge in comparison to the interview score, even if there were no false alarms. However, this may be the case for other written measures of vocabulary knowledge as well and thus warrants further investigation. A second innovative approach is taken by Stubbe and Stewart (2012), who compare existing scoring methods with a new one based on the standard least squares model used in linear regression. Findings using this formula in Stubbe (2013) indicate that the new regression formula shows slightly higher correlations with a translation test than the h-f formula, which again produced the highest correlations among the more conventional scoring formulae.

This brief overview shows that the problem of the best scoring method for Yes/No is still unsolved (Beeckmans et al. 2001; Meara 2010; Meara & Miralpeix 2017; Schmitt 2010). In many of the investigations above, the simple h-f formula produced similar results to more sophisticated correction formulae, but it has not yet been compared to the newest proposal, the logistic weighting function used in V_YesNo (for a comparison using the data of the present

---

[115] The use of timed Yes/No tests has since been further explored as a measure of lexical facility by Harrington (2018).

study, see section 6.4.1 as well as Table B.10 and Figure B.2 in Appendix B). In addition, and despite the use of scoring formulae, Yes/No tests cannot fully account for variation in relation to participants, for instance their confidence in responses or guessing strategies used (Eyckmans 2004; Eyckmans et al. 2007; Milton 2009). Furthermore, there is no evidence for participants' knowledge of the target words' meaning (Schmitt 2010) and thus the format automatically avoids the issue of multiple meanings (Beeckmans et al. 2001). The studies further point to issues with the test format such as instructions or length that have not received adequate attention in the past. The exact wording of instructions, the plausibility of pseudowords and the presence of cognates may affect test scores in ways that are yet unknown (Beeckmans et al. 2001). However, they also provide evidence for the validity of the format, most frequently in the form of concurrent validity with translation tests, the VLT or personal interviews.

The format also offers a number of advantages, the most important of which is the easy and quick administration. Time constraints are also one of the main reasons why in the end a Yes/No test was chosen as receptive vocabulary measure in the present study. In addition, the short time needed to take the test allows for a relatively high sampling rate which can increase reliability (Milton 2009; Schmitt 2010). Yes/No tests can easily be computerized, which is why some researchers including Nation (2013) have argued that they are particularly suitable as placement tests (see also Fairclough 2011; Harrington & Carey 2009). Moreover, the studies summarized above also provide some evidence for the validity of the format, most frequently in the form of concurrent validity. Finally, studies (Alderson 2005; Harsch & Hartig 2016) have shown that it correlates suitably well with other measures of language proficiency.

The test used in the present study, V_YesNo (Meara 2015a), is based on previous work by Meara, most notably the EVST by Meara and Jones (1990). As mentioned above, it consists of 200 items divided equally into 100 target words and 100 pseudowords and is scored using a new formula based on a logistic weighting function (Meara & Miralpeix 2017).[116] A computerized version is available from Paul Meara's homepage (http://www.lognostics.co.uk/), but for this project a paper-and-pencil version was constructed (see Figure 5.6), which also includes 20 translation items to check the reliability of students' self-report data. For the exact instructions used, please see the test included in Appendix A and the description on page 178. For details of test administration and data entry and scoring, please see sections 5.3.4.1 and 5.3.5.3.

The productive measure used in the present study is Lex30 developed by Meara and Fitzpatrick (2000). The test uses a word association task with 30 stimulus words to elicit a set of response words from participants.[117] Lex30 thus shares characteristics with both free and context-limited productive measures because there is "no predetermined set of response target words [..., but] the stimulus words tend to impose some constraints on the responses" (Meara & Fitzpatrick 2000: 22). The small sample of words elicited from participants is then analysed according to

---

[116] Further approaches to scoring the V_YesNo data are explored in sections 5.3.5.3 and 6.4.1.
[117] It is important to note that although a word association task is used in Lex30, it is not a word association test (Fitzpatrick & Clenton 2010).

frequency with a higher proportion of low-frequency words indicating greater productive vocabulary knowledge (see section 3.2.3).

In terms of construct, Fitzpatrick and Meara (2004) argue that Lex30 clearly elicits productive vocabulary knowledge because participants produce their responses in written form. They concede, however, that "subjects' knowledge of the words they produce could vary widely" (Fitzpatrick & Meara 2004: 71) from early stages of word knowledge to deeper knowledge involving meaning, collocations, associations and more (see section 3.1.1). At the minimum, test takers need to know the written form of the word in order to produce it on the test, but because of the timed association task (see below), it is likely that they have deeper knowledge of most words they produce. In an analysis comparing the constructs of Lex30, the PVLT (Laufer & Nation 1999) and an L1-L2 translation test with the help of Nation's (1990) taxonomy of word knowledge, Fitzpatrick (2007) asserts that Lex30 taps into productive knowledge of written form, productive knowledge of meaning concept and productive knowledge of meaning associations. The last point has, however, been contested by Kremmel (2017: 56), who argues that the test's format and instructions are more in line with receptive knowledge of meaning associations. In addition, he criticizes that Lex30's overall construct and the purpose of the test remain unclear, which affects the interpretation of scores. A second aspect that has been criticized in relation to construct is the fact that Lex30 only requires minimal productive knowledge and assesses recall rather than use (Read 2000). Still, as we have seen in section 3.2.3, the choice of productive vocabulary size tests is limited and most other available options such as the PVLT or L1-L2 translations tests also test form recall.

As mentioned above, the core of the Lex30 test is a list of 30 stimulus or cue words. Care was taken in the selection of these words, which had to fit three criteria: first, cue words have to be highly frequent so that they are well known to test takers. Second, cue words should not elicit "a single, dominant primary response" (Meara & Fitzpatrick 2000: 22), or in other words, they should not have strong, typical associations so that participants will produce a range of different responses. Third, the cue words should "generate[...] responses which are not common words" (Meara & Fitzpatrick 2000: 23), meaning that stimuli should provide test takers with a good opportunity to produce infrequent words. To operationalize these criteria, Meara and Fitzpatrick (2000) only included cue words from a list of the 1,000 most frequent words in English (Nation 1984), excluded all cue words for which one typical response accounted for more than 25% of all reported responses in the Edinburgh Associative Thesaurus (Kiss, Armstrong & Milroy 1973) and used L1 English speakers to check whether at least half of the responses given were beyond the 1,000 word level of Nation's (1984) list.[118] The final list of 30 cue words (see Appendix A) was presented in Meara and Fitzpatrick (2000) and can also be accessed via the online version of the test available at http://www.lognostics.co.uk/tools/Lex30/.

---

[118] Although the norms in the Edinburgh Association Thesaurus are based on L1 user data, Fitzpatrick and Clenton (2010: 539) argue that "the response features relevant to Lex30 – variety and frequency – tend to be cue-item specific, and hold true for native and non-native speaker responses equally".

In the Lex30 test test takers are provided with this list of cue words and are asked to write down the first four words that come to their mind as a response (see Figure 5.5). The test instructions given in Fitzpatrick (2007: 120) and Fitzpatrick and Clenton (2010: 548) read:

> Look at the words below. Next to each word, write down any other words that it makes you think of. Write down as many as you can (4, if possible). It doesn't matter if the connections between the word and your words are not obvious; simply write down words as you think of them.

In contrast, Meara (2009: 146) provides a slightly different set of instructions:

> In this test, you will see a list of 30 English words. Each word will make you think of several other words in English. Write these words in the boxes alongside each word.

In addition, these instructions include an example of a stimulus word and possible responses. The instructions used in the present study combine both of these suggestions and make students aware that only English words are acceptable as responses and that the administration of the test is timed. Translated into English, the instructions were

> In this vocabulary measure you will see a list of 30 English words. Write the first other English words that come to your mind next to each word. Write down as many as you can, if possible four words (there are four empty boxes next to each word). It doesn't matter what kind of connection there is between the first word and the other words, just write everything that comes to your mind down in English. For instance, someone has written these words for the word "animal": *elephant, farm, wild, feed*. Don't think about it too long, you only have 15 minutes overall!

Limiting the amount of time available is recommended, for instance, by Meara (2009: 134): "[Y]ou should encourage test-takers to work as quickly as possible, and data produced by test-takers who take a very long time on the test should be treated as unreliable." This is because the test should elicit spontaneous responses, presumably because these are expected to reflect words that the participants know well enough to fluently retrieve them from memory. In the studies by Fitzpatrick and colleagues (Fitzpatrick 2012; Fitzpatrick & Clenton 2010; Fitzpatrick & Clenton 2017; Meara & Fitzpatrick 2000) test takers were given 30 seconds per item; the test thus took 15 minutes overall. Walters (2012), who allowed as much time as participants needed, found that completion took between 15 and 30 minutes.

After data collection the samples elicited by the word association task are analysed in relation to frequency bands, similar to frequency-based lexical sophistication in free production measures (see sections 3.2.3 and 5.3.5.4). In line with the basic premise of the method, each word beyond the 1,000 word level of the frequency list scores 1 point, whereas the first 1,000 words do not receive any points. Since test takers produce a maximum of 120 response words, the maximum score theoretically is 120 points. Some researchers have raised concerns that this rather simplistic scoring method is problematic because it treats all response words beyond the 1K frequency level the same. For instance, Walters (2012: 184) argues that "a single Lex30 score might represent a variety of vocabulary profiles" and suggests presenting the results as lexical frequency profiles instead. In contrast, Fitzpatrick and Clenton (2010: 548) maintain that a more fine-grained distinction is not useful because of the small sample size of 120 words and because frequency lists "differ considerably in their assignment of words to lower frequency bands,

[while] there is a good deal of agreement over which words are the most highly frequent". In the original presentation of the Lex30 test Meara and Fitzpatrick (2000: 26–27) justify their scoring system in the following way:

> The lenient scoring method adopted for Lex30 – basically any slightly unusual word produced by the testee counts towards their score – means that testees are given credit at every possible opportunity. This contrasts sharply with the scoring practices typically used in more strictly controlled productive tests, where only the 'correct' response is counted. In Lex30, the stimulus word POTATO might cause a medical student to respond with CARBOHYDRATE, and a waiter to respond with MASHED. Both responses are 'unusual' in the sense that we are using that term, and so both are awarded a point. *In this way, we do not penalise students whose experience of words is influenced by special circumstances or special experience* [emphasis added].

One could argue that this scoring method is thus particularly appropriate for the present study as some of the vocabulary known by participants is likely to be influenced by the special circumstances or experiences of EE. In addition, for many purposes a single score can be more useful than a frequency profile.

A second point of criticism voiced by Walters (2012) is, however, much more pertinent and concerns the interpretation of scores: the numerical score produced by Lex30 does not correspond to an estimate of productive vocabulary size in terms of the number of words a participant knows productively. It is a score based on the number of infrequent words produced and as such it can be used "as a way of comparing individuals in terms of their breadth of vocabulary knowledge. [...] However, there is still the question of what a particular score means in terms of vocabulary knowledge" (Walters 2012: 184). This certainly is a critical issue, which also relates to Kremmel's (2017) criticism of Lex30's construct. The test developers are clearly aware of such concerns and Meara (2009: 136–137) highlights

> that Lex30 does not attempt to provide an accurate measure of the total productive vocabulary that the test-takers have at their disposal. It produces a score which we think might be related to this total, but should be treated with appropriate caution. The scores are probably reliable enough to allow for comparisons between groups.

Based on validation studies, Meara (2009: 136) states that "[a] good native speaker score is about 60 points" and Fitzpatrick and Clenton (2010: 539) report that "none of the native or non-native speakers we tested scored higher than 70, and most learner scores tend to be in the 10–40 range." In order to link these scores to more meaningful interpretations clearly much further research is needed (Walters 2012), but since the main goal of the present study is to compare participants among each other, the current scoring procedure is acceptable.

In addition to first impressions of a common range of scores among L1 and L2 speakers, validation studies of Lex30 have also produced other rather promising results. Meara and Fitzpatrick (2000) tried out the test with 46 adult EFL learners from different L1 backgrounds and found a large and significant correlation of $r = .841$ ($p < .01$) with a Yes/No test (Meara & Jones 1990). Based on this plausible relation between the receptive and productive vocabulary size of their participants they conclude that "Lex30 is sensitive to gross differences in vocabulary knowledge" (Meara & Fitzpatrick 2000: 26). Fitzpatrick and Meara (2004) also found that Lex30

was able to differentiate between L1 and L2 users of English and that there were significant correlations of moderate strength between Lex30 and the VLT and an L1-L2 translation test. The study by Walters (2012) compared Lex30 scores in three groups of Turkish learners (*N* = 87) at different proficiency levels and found that the mean test scores were significantly different between the three groups although there was some overlap. In addition, she also administered the PVLT and an L1-L2 translation task to the same three groups and found broadly similar results to Fitzpatrick and Meara (2004), although the correlations between Lex30 and the other two test forms were slightly higher in her study. Finally, Fitzpatrick and Clenton (2017) compared Lex30 to the Lexical Frequency Profile (LFP, Laufer & Nation 1995) in a study with 80 participants and found that their scores were not significantly related. This is surprising because both measures analyse text samples according to frequency, although Lex30 uses a word association task to elicit the sample, whereas the LFP employs essay questions. After two further experiments in which Lex30 is compared to newly designed tasks, the researchers then concluded that the Lex30 and the LFP differ in their "capture zone" (Fitzpatrick & Clenton 2017), but see Kremmel (2017) for a critical discussion.

Concerning reliability, Meara and Fitzpatrick's (2000) results indicate that the test has good internal consistency with split-half reliability showing a correlation of r =.84. Fitzpatrick and Meara (2004) investigated reliability with a test-retest approach: 16 L2 learners of English took the Lex30 test twice with a three-day gap. Results show that although only about half of the response words produced the first time were also produced the second time, the number of infrequent words produced remained roughly the same and the correlation was large with r = .866 (p < .01). Fitzpatrick and Clenton (2010) provide further evidence on reliability by exploring the results of two parallel test versions, which showed very similar scores with a medium correlation of r = .692 (p < .01) and no statistically significant differences. In addition, a computation of Cronbach's alpha based on the data of 35 test takers indicates that internal consistency is acceptable with α = .866. Fitzpatrick and Clenton (2010) also investigated the influence of the mode of elicitation on the results by administering the conventional written version as well as a spoken version of Lex30 to 40 Asian university students. Again, the results did not show any statistically significant differences, but the correlation was relatively weak with r = .39. Fitzpatrick and Clenton (2010: 546–547) take these results to mean that "the tests might not work in exactly the same way, or indeed that test takers' oral ability might not match their written ability". In addition to these validation studies, Lex30 has also been taken up quickly by other researchers and used in a number of studies, particularly in Spain and in relation to CLIL contexts (Alejo González & Piquer Píriz 2016; Jiménez-Catalán & Moreno Espinosa 2005; Moreno Espinosa 2010). It has also been used to track changes in the productive lexicon in a longitudinal case study (Fitzpatrick 2012) and to investigate the relationship between meaning-focused listening input and productive vocabulary knowledge (Noughabi 2017).

In sum, Lex30 presents an innovative approach to the measurement of written productive vocabulary knowledge as it combines characteristics of free production tasks and controlled, context-free tests. Although it is still an "exploratory and experimental" (Meara 2009: 132) measure that is clearly in need of further validation, it has a number of advantages. First, it uses an easy to explain task with a high level of face validity that can be administered in a relatively short amount of time (Meara 2009; Meara & Fitzpatrick 2000). Second, unlike other measures such as the PVLT, Lex30 requires very little receptive vocabulary knowledge for the elicitation of samples of productive vocabulary (Fitzpatrick & Clenton 2010; Kremmel 2017). Third, the data can potentially be analysed with other word lists and for other research purposes as suggested by Fitzpatrick and Clenton (2010), an option that is explored in the present study (see section 6.4.7). However, the test is clearly in need of further validation and more research is needed to clarify issues of construct definition and score interpretation (Kremmel 2017; Walters 2012). In the end, Lex30 is not a perfect test, but it is a useful research tool and until a better, thoroughly validated measure of productive vocabulary becomes available, it remains a good and viable choice.

This section provided detailed information on the development and selection of all four instruments used in the quantitative strand: the Extramural English Questionnaire (EEQ), the Extramural English Online Language Diary (EEOLD), and the two vocabulary measure V_YesNo and Lex30. It also presented the content and format of the final versions used in the study, the original German-language versions of which can be found in Appendix A. In the following section, details are given on how these instruments were used to collect data from the participants.

5.3.4 Quantitative data collection

This section provides details about the data collection procedures used in the quantitative strand. The overall procedure for quantitative data collection has already been described in section 5.2.3: in the first session, the EEQ was administered (see section 5.3.4.1) and subsequently the instructions for the EEOLD were provided (see section 5.3.4.2). In the second quantitative data collection session the two selected vocabulary measures, Lex30 and V_YesNo, were completed by the students (see section 5.3.4.1). Originally three data collection sessions in three lessons had been planned, but it quickly became clear that most teachers and heads of school would not allow the use of that much class time. Hence, instruments and data collection procedures were adapted to fit a time frame of two 50-minute lessons to suit the needs of the school research context.

To link the quantitative instruments administered in different sessions and online to one participating student, an anonymous code was used: each instrument contained a set of instructions to construct a five-digit code following the example of Kearney et al. (1984) as cited

in Dörnyei (2010: 81–82).[119] Figure 5.9 presents the instructions for the code translated into English (see Appendix A for the original version):

My code:	☐☐	2nd and 3rd letter of your first name
	☐	last letter of your mother's first name
	☐☐	month in which you were born (e.g. 03 for March)

Figure 5.9: Instructions for the anonymized code to link an individual participant's instruments

In line with the suggestions found in Dörnyei (2010: 81–82) the code thus consisted of "specific code elements that are well known to them [the participants] but not to the researchers". Participants simply filled in the boxes on the left and thus generated their own pseudonym for the study. In addition, I then added the class code (e.g. SA01) to the five-digit code to avoid confusion of identical codes in different classes. When explaining this procedure to the students, it was important to highlight that while this code was in theory not completely anonymous, I did not have access to the information, such as their birth dates or their mothers' names, that would allow me to trace any information back to them and they could therefore rest assured that their responses would be confidential.

5.3.4.1 Questionnaire and test administration

This section describes the data collection procedures for the three offline instruments: the EEQ and the two vocabulary measures, V_YesNo and Lex30. They were all administered in paper-and-pencil format because in many Austrian schools computer rooms are not easily available or do not have enough machines for a whole class.

As described above, data collection with the EEQ and the two vocabulary tests took place in two 50-minute sessions on different days. The teachers were not usually present during these sessions to emphasize the confidentiality of the students' responses. It was planned that as the researcher I would administer all instruments myself; however, it soon became obvious that that would not be possible for scheduling reasons: in two schools (SC and SF) English groups are split across classes and therefore they always take place in the same slot of the timetable. While it would have been possible for me to collect data in these groups on different dates, the teachers preferred concurrent data collection sessions in both groups. Fortunately, with Magdalena Hahn, who was collecting data for her study at approximately the same time, a colleague familiar with all instruments and data collection procedures agreed to help in these instances.[120] Her support enabled data collection in schools SC and SF to go forward as planned, as her presence ensured that established procedures were followed. In addition, she also supported data collection in one session in school SA; thus, overall ten out of 24 data collection sessions were carried out in parallel (for an overview of all dates see Table 5.2 in section 5.2.3.).

---

[119] Researchers conducting school-based research in Vienna need to ensure the anonymity of participants and are not allowed to link data directly to participants' names.
[120] As discussed in section 3.3.1 and 4.4, Magdalena Hahn conducted a partial replication of the present study with students in vocational business middle schools (Hahn 2017) and used the same paper-and-pencil instruments.

At the beginning of the first data collection session, I checked whether all students present had parental permission to take part in the study and handed out questionnaires. After reading through the introduction and the instructions together, the students filled in their codes. In addition, I asked them for a definition of leisure time and explained that I defined it as all time not spent in lessons; thus, breaks at school were considered leisure time, but having to visit a great aunt with their parents counted as well, although the students might not have chosen to do so. Furthermore, we also briefly discussed what was meant by English-language music, films or series because in the pilot studies some students had misunderstood this term as originating from an English-speaking country rather than as being listened to or watched in English as the original language. Students generally took between 25 and 40 minutes to fill in the EEQ. Once a student had finished I collected their questionnaire and traded it for a sweet of their choice as a small token of gratitude. After all students had finished, I handed out the instructions for the EEOLD (see section 5.3.4.2).

In the second data collection session I briefly introduced this day's programme and highlighted importance of completing the two vocabulary measures on their own. If there were not enough desks for each student to have one on their own, I asked teachers to provide me with partition walls that are usually used for tests and distributed these among the students who shared a desk. Then I handed out Lex30, we filled in the code and read through the instructions together. I emphasized that there were no wrong answers on this measure as long as they wrote down English words and explained that they should write down the first words that came to their minds because there was a time limit. They had five minutes for each page, if they were done earlier they could go to the next page, but when I sounded the bell after five minutes, everybody had to turn to the next page.

Because of this time limit, the time needed to complete Lex30 always was 15 minutes. I then collected the tests and students could again select a sweet as a small favour. If we had started the session punctually, there was time to fill in the EEOLD for the day before in between the two vocabulary measures (see section 5.3.4.2). I then distributed V_YesNo and again we filled in the code and went through the instructions. I explicitly told students that points would be taken off their score, if they ticked "Yes" for a pseudoword that does not really exist in English and asked them to be careful. They were also told that after the checklist items, there were items to check their actual knowledge for some words and that they were not allowed to turn back. Time needed to complete V_YesNo was usually between four and eight minutes, but never more than ten. The students then completed the translation task, in which they were asked to translate 20 items into German or provide an English synonym or explanation.[121] Once they were finished, I asked them to fill in the short post-test questionnaire on the last page.

---

[121] The test takers were made aware that the translation task did not contain any pseudowords. In order to ensure that this information did not impact their result on the V_YesNo test, care was taken during the administration to prevents participants' from going back to the V_YesNo test in the first part of the test booklet once they had seen the translation items.

At the end of the session I thanked the class for their participation and informed them that I would come back in spring with their results, if they wanted to look at them. It was emphasized again that only they as a group would receive them, not their teachers. Furthermore, I announced that in spring I would ask if some of them wanted to take part in a voluntary group interview to discuss some of the results among other things. After the last data collection session, the teachers also received a small token of gratitude for their support. Overall, the data collections sessions proceeded without any major issues. Some participants missed one of the two sessions due to absences (see sections 5.3.5.1 and 5.3.5.3), but otherwise there were no problems which could have influenced the quality or comparability of the collected data.

5.3.4.2 Online data collection

In the present study only one instrument was administered online: the Extramural English Online Language Diary (EEOLD). Since the language diary was meant to be filled in every day over a period of one week, online data collection was the preferred option because students could easily access the diary whenever they wanted to and there was no need to involve teachers in the daily data collection process. However, online data collection also holds potential pitfalls; some of which became apparent in the present project.

First, data protection is an issue when collecting data online, since it is not always clear where the collected data are stored. Preparation for my online data collection happened to commence shortly after the Safe Harbour Privacy Principles were overturned by the European Court of Justice, which affected the possibility to legally store data on servers in the United States of America (see for instance, Monteleone and Puccios's 2017 analysis for the European Parliament). Since Qualtrics, the platform I intended to use, is based in the US, I first needed to ensure that the data would be stored on European servers, which fortunately was the case according to the Qualtrics Security White Paper (Qualtrics LLC 2015). In addition, no personal data were collected in the EEOLD since all the sociodemographic information needed was included in the EEQ (see section 5.3.3.1). Hence, online data collection using the EEOLD could go forward and was piloted in the full quantitative pilot study (see section 5.2.3).

In the main study students received the instructions after they had completed the EEQ in the first data collection session. The instruction sheet for the EEOLD (see Appendix A) gives reasons for why they should fill it in regularly and information on what to expect. In addition, an abbreviated link as well as a QR code to access the online survey are provided. In most sessions there was time to go through the instructions together and to stress the importance of filling in the EEOLD every day. If possible, participants were asked to fill in the diary at the end of the day, or early the next day, so that their memory would still be accurate. To remind students about the language diary, I also put up A3-size coloured posters including the link and the QR code in their classrooms.

In order to connect the EEOLD to the other instruments used, the students first entered their code when accessing the diary (see section 5.3.3.2). However, while monitoring the online

activity for the first three classes in which the EEOLD was introduced, it became clear that identical codes could appear in different classes. This was not a problem with the paper-and-pencil instruments because I had added the class code, but it emerged as an issue in the online data collection. Fortunately, the two participants could be disambiguated based on the week in which they had filled in the diary, but the data collection procedure was adapted in all subsequent classes: the code field was changed to allow six instead of five characters and all classes were told to add the first letter of their school in addition to their code when entering it in the EEOLD. However, even with this measure some codes remained unclear, most probably because of typos. Some of these could be related to participants because of the week in which they filled in the EEOLD, but some diary entries could not be matched with any participant and thus had to be excluded from data analysis (see section 5.3.5.2).

At the beginning of the second data collection session I reminded all classes to fill in the EEOLD and asked whether they had experienced any problems. Those who had completed it usually had not, but many students had not done so regularly. Therefore, students were asked to fill in the language diary for the previous day once they had handed in Lex30 (see section 5.3.4.1), if they had not already done so. However, despite this measure low response rates remained an issue with the EEOLD, as is discussed in more detail in section 5.3.5.2.

5.3.5 Quantitative data preparation and descriptive analysis

After having collected the data, they needed to be scored in the case of the two vocabulary tests and all data needed to be computerized and prepared for data analysis. For this reason, detailed codebooks were prepared for the EEQ and the EEOLD and scoring protocols for Lex30 and V_YesNo. The compilation of these codebooks and scoring protocols already began during the analysis of the pilot data, but the documents were adapted and refined for the main study. After data entry was completed, the data were cleaned and checked for consistency and reliability before transformations were applied where necessary. In a first step of analysis descriptive statistics were computed for all variables in addition to graphical explorations of the data. In the following, details are provided on the procedures followed for each instrument.

5.3.5.1 Data entry and preparation for the EEQ

Due to administration during school lessons (see section 5.3.4.1) the response rate for the EEQ was very high, in the final sample of 201 participants only 12 students did not complete the questionnaire because they were absent on the day of data collection. After data collection, the data were computerized in accordance with the established codebook. Numeric data based on closed items were entered into SPSS (Version 22.0, IBM Corp. 2016) and responses to open-ended items which could be coded numerically were added as well (e.g. items 1e, 2c, 2d, 6a). Longer responses based on the remaining open items were collected in a separate Excel spreadsheet.

Subsequently, the numeric data were cleaned and prepared for further analysis in several steps. Frequency tables were inspected for all input variables to identify impossible values due to typos outside the pre-defined value range. To identify typos within the range of values, random spot

checks were made by entering one questionnaire per class in separate data sheet and comparing it to the originally entered data. Following the process of data cleaning and checking, several transformations had to be applied, such as recoding values for negatively-worded statements and calculating new variables. Some of these new variables were simple transformations such as the computation of age from the year of birth, but most were concerned with the establishment of higher-order variables to reduce the number of variables. For example, the information provided in EEQ item 6a on the languages spoken with different people at home was summarized in a variable called "number of home languages", which represents a simple count of different responses given in this item set. New variables were also derived from the list of 64 EE activities presented in item set 2a. For example, the two individual items "watching films on the internet (e.g. Netflix, …) with subtitles" (item 2a034) and "watching films on the internet (e.g. Netflix, …) without subtitles" (item 2a035) were combined to a summary variable "watching films online" using the maximum value of the two items. Thus, if a participant said they watch films on the internet with subtitles only a few times a year (= value 2), but without subtitles they watch them a few times per month (= value 3), this participant's score on the derived summary variable is 3 indicating that overall they watch films online at least a few times a month. In addition to creating summary variables for EE activities they were also recoded according to language skills for further analyses.

The data transformations in relation to the information on SES elicited in section 7 of the EEQ (see section 5.3.3.1) warrant special attention. Data provided on the level of parental education in item set 7c were coded according to the International Standard Classification of Education (ISCED, UNESCO Institute for Statistics 2012) as applicable to the Austrian education system (Bundesministerium für Bildung & Bundesministerium für Wissenschaft, Forschung und Wirtschaft 2017). Information on parental occupation was collected through open responses in items 7d to 7g, in which students were asked to indicate their parents' professions and give a brief description of their jobs. These data were then categorized according to the International Standard Classification of Occupations (ISCO-08, International Labour Office 2012) and later transformed into a score on the International Socio-Economic Index of occupational status (ISEI, Ganzeboom 2010) using SPSS syntax provided by Harry Ganzeboom on his website (Ganzeboom & Treiman 2010). After conversion of the classification codes into ISEI scores, the highest parental score per family was taken as an indicator of family SES.

Scores on the ISEI scale can range between 11.56 for the category 'Field crop and vegetable growers' and 88.96 for judges (Ganzeboom & Treiman 2010; Pham, Freunberger & Robitzsch 2014). Higher scores represent a higher social status resulting from a job with higher educational requirements and a larger income. While it is difficult to meaningfully interpret individual ISEI scores, they can be used to rank the students' families in terms of socioeconomic status in later analyses (see also PISA technical report, OECD 2014). In the present study 167 students answered the questions about parents' occupation in enough detail for a classification to be made about at

least one parent, while the remaining participants did not answer these questions or did not provide sufficient information.

After completing the data transformations, internal consistency was calculated for the multi-item scales tapping into attitude constructs (item set 3), the use of VLS (item set 5b) and the frequency of engagement with EE activities (item set 2a). Although such rating scales are sometimes treated as interval data in the social sciences (see Döring & Bortz 2016: 250–251 for a discussion), they are regarded as ordinal data in the present study because the response options used in the EEQ rating scales are not equidistant and therefore the data do not constitute an interval, but an ordinal scale (Dörnyei 2010: 92). For this reason, ordinal α, a reliability coefficient for ordinal item response data (Gadermann, Guhn & Zumbo 2012), was calculated rather than Cronbach's α. Interpretation of ordinal α can, however, follow the guidelines for Cronbach's α according to Gadermann, Guhn and Zumbo (2012), which generally report 0.7 as a cut-off point for reliability (e.g. Cohen, Manion & Morrison 2011: 640; Field, Miles & Field 2012: 799).

For the five subscales tapping into participants' attitudes towards English (see section 5.3.3.1), ordinal α was below the generally accepted threshold of 0.7, as shown in Table 5.7. This result is not wholly unanticipated due to the small number of three or four items per scale, but as a consequence, the attitude data will only be reported descriptively in Chapter 6 and not be used in any further analyses.

Scale name	Scale description	N of items	ordinal α	M	SD
Future	English is important for the future.	4	0.67	3.35	0.46
International	English is vital for international communication and travelling.	4	0.55	3.20	0.41
Youth	English is especially important for young people.	4[1]	0.50	3.10	0.43
Austria	English plays a significant role within Austria.	4	0.65	2.80	0.52
German	English is regarded as better than German.	3[1]	0.56	2.80	0.67

[1] One item was dropped due to lack of fit in each of these two scales.

Table 5.7: Reliability coefficients for attitude scales

For the VLS item set consisting of nine rating scales, ordinal α was 0.61 and thus cannot be regarded as a reliable scale either. Hence, the results on participants' use of vocabulary learning strategies are again described in Chapter 6, but are not used in inferential analyses. Finally, the calculation of ordinal α for the 65 items in the EE scale produced a result of 0.92, which means that this scale can be regarded as internally consistent and can be transformed into a summary variable. Since the data are regarded as ordinal, the median value was used instead of the mean to create an EE median score.

In a next step, descriptive statistical analyses were carried out for all variables (see section 5.3.6). Frequency distributions and relationships between variables were explored graphically. Moreover, the SPSS data file was prepared for further analysis in RStudio (Version 1.2.1335,

RStudio Team 2018), with which all statistical models were computed. In parallel, qualitative data from the open items were subjected to content analysis and grouped into categories (e.g. item set 1i, items 5c and 5d) to allow the establishment of frequency counts.

5.3.5.2 Data preparation for the EEOLD

After the data collection period for the EEOLD had closed, data were downloaded from the Qualtrics platform. The response rate was unfortunately rather low since the EEOLD had to be filled in during students' free time and was not part of any school assignments. In addition, no contact details could be collected to remind students about completing the EEOLD because of the strict privacy guidelines of the educational board of Vienna. As a result, only 485 diary entries were downloaded from the platform, which is far below the expected response rate. After removing partially completed entries with less than 50% progress 473 language diaries by 155 students remained. This means that out of the total number of 224 participants, 69 students did not fill in the diary at all and some only filled it in once or twice. Clearly, a lack of motivation to fill in online instruments which results in a low response rate is a major disadvantage of online data collection. In addition, several of these diary entries still had to be excluded: 56 entries were by participants who did not meet the criteria for inclusion in the sample (see section 5.3.2) and 7 entries could not be linked to individual participants as the codes given were non-existent among those given on the EEQ and vocabulary tests and no obvious mistakes (such as changing the order of two letters) could be identified. After these data cleaning procedures 410 diary entries by 130 students remained to be used.

The data were then processed further using SPSS. Several process variables used by Qualtrics were deleted and the remaining variables and values were renamed in accordance with the pre-established codebook. In addition, all time estimates were converted into minutes to allow for calculations. An estimate for time spent with EE was then computed separately for each language diary entry and, if more than one entry was available, a mean estimate was calculated for each participant. Moreover, estimates were also computed for each of the five sections tapping into listening, reading, gaming, writing and speaking.

A close inspection of the data showed that several estimates were incredibly high; indeed, the maximum amount of time spent with EE estimated by one participant on one day was 1685 minutes or 28 hours and 5 minutes. Clearly, no human being can make use of more than 24 hours a day, despite the fact that we all sometimes wish we could. A possible reason for such implausible estimates could be that students engage in EE activities simultaneously, for example, by using social media while listening to music, and included these times twice in the diary. Still, even if multi-tasking may play a part, any estimates higher than 720 minutes or 12 hours are highly improbable considering that students also need to sleep and go to school. 12 hours were therefore set as a threshold and any diary entries containing higher estimates for one day were removed. As a result of this data cleaning procedure another 27 EEOLD entries were excluded, leaving a final sample of 383 diary entries by 118 participants to be used in the analysis.

5.3.5.3 Scoring V_yesno

Like for the EEQ, the response rate for both vocabulary tests was high because of the administration during school lessons. A total number of 198 students completed the receptive vocabulary test as 26 students were absent in this data collection session. Furthermore, 19 students did not meet the criteria for inclusion (see section 5.3.2) and four participants did not fill in the paper-and-pencil test completely and therefore had to be discarded.[122] As a result, 175 participants' V_YesNo tests could be used for further analysis.

The scoring protocol for V_YesNo closely followed the procedure described in Meara and Miralpeix (2017). As described in section 5.3.3.3, scoring of Yes/No tests is based only on participants' 'Yes' responses, which either fall into the category of a hit or a false alarm. V_YesNo uses a new scoring formula based on an S-shaped logistic weighting function, which "avoids excessive penalisation for guessing, and is more generous towards guessing if the test taker gets most of the real words correct" (Meara & Miralpeix 2017: 119). This means that this latest scoring formula does not only take the number of FAs in a given segment into account, but also the number of hits. Hence, for the same number of FAs, the score of a test-taker who got a greater number of hits is adjusted less than the score of a someone who had fewer hits. As Meara and Miralpeix (2017: 119) state "[t]his approach is intuitively correct" and also avoids the problem of negative scores, which was an issue with earlier formulae such as Δm (see section 5.3.3.3).

To calculate the score using the S-shaped logistic weighting function, the 200 items of V_YesNo are split into 10 blocks of equal length, each consisting of 10 target words and 10 pseudowords. For each block, the number of hits and FAs is counted and adjusted by a correction factor A. The equation for the correction factor A given in Meara and Miralpeix (2017: 120) is:

$$A = 1 - \left(\frac{w(f)}{w(h)}\right)$$

with $w(f)$ and $w(h)$, the weighted number of false alarms and hits, being taken from the following table:

x	0	1	2	3	4	5	6	7	8	9	10
w(x)	0	1	3	6	10	15	20	24	27	29	30

Table 5.8: Values for w(f) and w(h) taken from Meara and Miralpeix (2017: 120)

To obtain the estimate of vocabulary size for one of the 10 blocks, the number of raw hits is multiplied by the correction factor A and then multiplied by 100, since the 10 target words of one block represent knowledge of 1000 words:[123]

$Vsize = h * A * 100$

---

[122] Four of the 23 students who fell into one of the categories for exclusion were not present at school on the day of data collection.
[123] This is the case because in V_YesNo a sample of 100 target words represents the 10,000 most frequent words of English (Meara & Miralpeix 2017: 120).

The overall estimate of vocabulary size then is the sum of the vocabulary size estimates of all 10 blocks. The complete formula can thus be summarized as:

$$\text{Vsize} = \Sigma \left(h_s * 100 * \left(1 - \frac{w(f_s)}{w(h_s)}\right)\right)$$

In practice, the 10 blocks used for scoring were marked on the paper-and-pencil version of V_YesNo and each block was scored manually using a green maker to indicate hits and an orange marker to highlight false alarms. Next to each block the number of hits and FAs was recorded and then entered into a prepared Excel spreadsheet. Again, spot checks were made by scoring and entering several tests twice to highlight any problem with data entry. In a next step, the corresponding weighted values *w(f)* and *w(h)* taken from Table 5.8 were inserted manually. The number of total hits, total false alarms and the vocabulary size estimate were then calculated automatically for each participant by Excel formulae.

In addition to adjusting the score downwards for overestimation with the correction formula, Paul Meara recommends discarding tests which show a large number of false alarms as unreliable because the results could be skewed despite the application of a correction formula. He suggests using 10 false alarms, i.e. 10% of the total number of pseudowords, as a reliability threshold (personal communication, 10.3.2016). In the present study the number of hits ranged from 18 to 86 out of 100 target words with the mean number of hits being 52.6 ([50.47, 54.92], SD = 14.81) and the median 52 ([48, 55], N = 175). Concerning false alarms, the minimum number was 0 and the maximum 41 with a mean false alarm rate of 7.77 ([6.79, 8.98], SD = 7.22) and a median of 5 ([5, 7], N = 175). Figure 5.10 summarizes the mean rates with regard to the stimulus response matrix. The mean FA rate is considerably lower than in some previous studies (e.g. Eyckmans 2004; Harrington & Carey 2009), but similar to results reported in Mochida and Harrington (2006). Reasons for the low FA rate could be the specific instructions used (see section 5.3.3.3), which followed suggestions by Mochida and Harrington (2006).

	Response alternative	
	Yes	No
Target word	*Hit* 0.53	*Miss* 0.47
Pseudo-word	*False alarm* 0.08	*Correct rejection* 0.92

Figure 5.10: Stimulus-response matrix for Yes/No tests including mean rates for the present study

Despite the relatively low false alarm rate, the maximum number of 41 false alarms indicates that test takers in this sample have exceeded Meara's proposed reliability threshold; in fact, 44 participants marked more than ten pseudowords as known. As dropping all of these from the analysis would have greatly diminished the amount of data to be analysed, a more lenient

reliability threshold of 15 false alarms, which would result in the exclusion of 25 participants, was also investigated. Table 5.9 displays descriptive summary statistics for all 175 completed tests, for the 150 tests which meet a lenient reliability threshold of 15 FAs, and for 131 tests that were below the strict reliability threshold of 10 FAs. As can be seen the minimum and maximum score are the same for all three reliability categories, but the mean and median scores across all tests (i.e. including tests with more than 15 FAs), are notably higher than those of the other two groups, which suggests that the advice to exclude high FA rates as a symptom of overestimation that cannot fully be taken into account by a correction formula is justified.

	N	*Min*	*Max*	*Mdn*	95% CI	*M*	95% CI	*SD*
all tests	175	1600	8447	4648	[4458, 4969]	4831	[4630, 5048]	1410.77
lenient (<15 FAs)	150	1600	8447	4564	[4314, 4989]	4809	[4578, 5053]	1493.85
strict (<10 FAs)	131	1600	8447	4521	[4203, 4903]	4787	[4527, 5058]	1529.37

Table 5.9: Summary statistics for a lenient and a strict reliability threshold as well as the complete sample of V_YesNo tests

In light of the considerable reduction of data if the strict reliability criterion were applied, I decided to exclude only tests which exceeded the reliability threshold of 15 FAs; thus, leaving 150 tests for further analysis. This decision to use the more lenient reliability threshold is justified here because the present study combines the use of a scoring formula with a maximal number of FAs, unlike other studies such as Schmitt, Jiang and Grabe (2011), who only relied on the reliability threshold. In addition to applying this reliability threshold, the reliability of the instrument in the present context was also explored by calculating Cronbach's α as a measure of internal consistency. Results indicate that reliability was sufficiently high with α = .89 [.87; .91].

Finally, the translation task included with V_YesNo (see section 5.3.3.3) can also be used to investigate whether scores on the Yes/No test accurately reflect students' knowledge of the form-meaning link. As the construct of the test is regarded as meaning recall (see sections 3.2.3 and 5.3.3.3), it can be compared to the translation items, which also tap into meaning recall. Ideally, participants' responses on the two measures correspond and indicate a correct judgement; i.e. target items that were selected as known on V_YesNo are translated or explained adequately, and target items that were not selected on V_YesNo are translated or explained incorrectly or not at all. Mismatches between the information provided on the V_YesNo test and in the translation task result in incorrect judgements, showing that participants either overestimated or underestimated their knowledge on the V_YesNo test.

20 target words, thus one fifth of the total number of target items included in V_YesNo, were chosen as translation items to check participants' actual knowledge of these words (see also section 5.3.3.3). Care was taken to include nouns, verbs and adjectives as target items and participants were asked to translate these into German or explain their meaning in English. The

translations items were originally included to deter participants from guessing. During data analysis, however, it became clear that students' performance on the 20 translation items could be used to investigate the validity of the V_YesNo data and to score the test data in a different way.[124] In practice, the translation tasks of the 175 students whose V_YesNo tests could be included in data analysis were scored manually. However, one participant did not fill in the translation task; hence, the final number of tasks to be analysed was 174. An overview of the accepted translations and L2 explanations can be found in Appendix A (Table A.1). In general, translations or explanations that demonstrated some knowledge of the form-meaning link or minor misspellings were accepted as correct, thus giving participants credit for partial knowledge of word meaning.[125] After the translation items had been scored the information whether a target word was selected as known on the V_YesNo test and whether it was translated or explained correctly were entered into an Excel spreadsheet. A logical formula was used to determine whether the information of the two measures corresponded and thus represented a correct judgement. [126]

Since the total number of items on the translation was 20, the number of correct judgements divided by 20 gives the proportion of correct judgements, which was calculated individually for each of the 174 students. On average the proportion of correct judgements was 74.08% ([72.19%, 75.80%]) for all 174 participants, meaning that in nearly three quarters of the cases the participants' choice of response on the V_VesNo test corresponds to (a lack of) knowledge demonstrated on the translation task. As Meara (personal communication, 10.3.2016) argues that V_YesNo test performances with high false alarm rates are unreliable, it is interesting to see whether the ratio of correct judgements is different for those participants who produced fewer than 15 false alarms. Unexpectedly, however, it is very similar and only slightly higher with a mean of 75.94% ([74.12%, 77.60%], N = 149).

The findings for receptive vocabulary size presented in Chapter 6 first explore the results for the V_YesNo test in greater detail by investigating methodological issues, in particular the use of different scoring formulae and the possibility to use the proportion of correct judgements as a correction factor for the raw number of hits ($h \times CJ\%$) instead of the S-shaped logistic weighting formula. Subsequently, the relationship between the V_YesNo scores, engagement with extramural English and other influencing factors included in the design of the present study is explored.

---

[124] I am grateful to Benjamin Kremmel, Tineke Brunfaut and in particular Norbert Schmitt for their suggestions for a scoring method based on the translation items.

[125] A good example for partial knowledge is the translation of *elegance* as German *elegant*, which disregards part of speech. Such translations were generally accepted, except for the English verb *analyse*, which has the same orthography as the German noun *Analyse*. For this reason, students had to demonstrate knowledge of part of speech, too, for this item by giving the corresponding German verb *analysieren*.

[126] It is important to note that a correct judgement is formed both by a "Yes" reponse on the V_YesNo test and a corresponding correct translation or explanation as well as by a "No" response and a wrong or missing translation or explanation. In both cases the judgement made by the participants about their knowledge of a specific items was accurate.

5.3.5.4   Scoring Lex30

As mentioned in the previous section, both vocabulary tests were completed by 198 students, but tests by 19 participants had to be discarded due to the exclusion criteria presented in section 5.3.2. Additionally, the Lex30 tests of seven participants could not be used as they filled in less than half of the response words or missed a complete page.[127] Hence, 172 Lex30 tests remained to be used for further analysis.

Again, a detailed scoring protocol was used; it was based on recommendations in Meara and Fitzpatrick (2000) and Meara (2009) as well as on information obtained directly from Tess Fitzpatrick and Tom Caton, a fellow PhD student (personal communications, 29.4. and 10.6. 2016). The scoring process entails entering and cleaning the data before comparing it to a frequency list; in this case the JACET 8000 word list (Ishikawa et al. 2003; Uemura & Ishikawa 2004) was used as recommended by Meara (2009: 133).

To begin the scoring process, the data produced by participants were entered into individual excel spreadsheets named with participant codes. At this point, an excel file contained all tokens produced by a student including misspellings and proper names. Next, the raw data were colour-coded to differentiate between categories that needed to be treated differently in the scoring process. These categories are:

a)   non-English words and non-identifiable misspellings
b)   proper nouns, names and acronyms
c)   misspellings of identifiable English words
d)   inflected forms
e)   British English spellings[128]
f)   multiword units and phrasal verbs
g)   cue words

Following this colour-coding procedure, the data were cleaned and the scoring words were entered in a second column next to the original answers: for misspellings the word form was corrected and British spellings were changed to American forms. Inflected words forms such as conjugated verbs or plurals were lemmatized according to the instructions by Meara and Fitzpatrick (2000), which are based on Bauer and Nation's (1993) criteria.[129] However, these criteria are not always in accordance with the JACET 8000 list now recommended for scoring: first, the -er suffix was found to be problematic because words such as *farmer* and *farm* or *gardener* and *garden* have separate entries in the JACET 8000 list; hence, lemmatizing these

---

[127] In Lex30 test takers are asked to write down four response for 30 cue words resulting in a maximum number of 120 words per sample (see section 5.3.3.3). It is accepted that test takers may not always be able to think of four responses to any given cue; however, samples with less than 60 response words overall were excluded as unreliable as the word count directly affects the score and makes comparisons with scores based on larger samples less meaningful.

[128] British spellings variants are colour-coded separately and changed to American forms because the word list used for scoring, JACET 8000, uses American spelling.

[129] Lemmatization criteria given in Meara and Fitzpatrick (2000) use level 2 and 3 of Bauer and Nation's (1993) classification. Level 2 includes inflectional suffixes such as 3rd person singular forms, past forms and gerunds, as well as comparative and superlative forms of adjectives, and possessives. Level 3 contains the most frequent regular derivational affixes such as *-er, -ish. -ly, -ness* or *un-*.

forms seems unnecessary. Even more problematic is the response word *computer*, which according to the criteria set out in Meara and Fitzpatrick (2000) should be lemmatized to *compute*. However, while *computer* is a highly frequent 1K word in the JACET 8000 list, *compute* is not, and a participant would therefore receive a point for the lemmatized version, but not for the word form they originally produced. After consulting with Tess Fitzpatrick (personal communication, 10.6.2016), I therefore decided to check the JACET 8000 list in cases of doubt and not to lemmatize forms included in the list.[130]

Another category that had to be cleaned were multiword units, since at the moment they cannot be scored using established frequency lists (see section 3.1.1). For multiword units and compound nouns the frequency of their constitute parts was checked: if one part was outside the 1K band of the JACET list, that word was used for scoring; if all parts were within the 1K band, the first one was retained unless it was a repetition of a previous response or a cue word. This approach is in line with Meara (2009: 134), who emphasizes that multiword units are acceptable, but adds the following comment regarding scoring:

> You will also need to decide how you are going to handle multi-word responses. Lex30 will generally not recognise these responses, which fortunately tend not to appear very often. The best approach is to simplify these items and score only the least frequent of the words they contain. E.g., If [sic] a test-taker responds to the stimulus word **attack** with **death or glory**, **glory** would count as a scoring word in its own right (Meara 2009: 138, emphasis in original).

A particular issue in this regard are compound nouns containing cue words: as highlighted by Fitzpatrick (personal communication, 10.6.2016), "the repeated cue word should, by definition, be a non-scoring (1k) word". However, she recognizes a problem first pointed out by Jiménez Catalán and Moreno Espinosa (2005: 37): changing the frequency list used for scoring from the word list originally used for selecting the stimulus words (Nation 1984) to the JACET 8000 list (Uemura & Ishikawa 2004) resulted in an inconsistency because not all cue words are included in the 1K band of the newer JACET 8000 word list (e.g. *attack, experience, hope*). However, in line with Fitzpatrick's statement above and because participants could simply copy cue words as their responses, I decided not to award any points to stimulus words used as responses. Consequently, if a multiword unit or a compound contained a cue word, such as *tooth* in *gold tooth* or *trade* in *fair trade*, the other part of this lexical unit was retained.

Once this data cleaning procedure was completed, the corrected list was copied to a new tab in the Excel file. This list, which no longer contained any misspellings, proper nouns, acronyms, inflected forms or multiword units, included all potential scoring words. The final list was then copied to a text file (.txt) so it could be read by the programme AntWordProfiler (Version 1.4.1w, Anthony 2013). Using this programme, the samples produced by participants were first compared to a list of the Lex30 cue words to double-check that these had been removed. The

---

[130] For this reason, word forms including the *-er* suffix, such as *computer, farmer, gardener* or *sweater* were not lemmatized if they were included in the JACET 8000 list. The same is true for *boring* (1K) and *crazy* (1K), which in their lemmatized forms (*bore* and *craze)* would also score points. In addition, mass nouns such as *glasses, pants* or *jeans* were not regarded as plural forms and thus were not lemmatized either.

sample of words produced by a given participant was then compared to the eight levels of the JACET 8000 list. The output of the programme sorts all response words in the text file according to the frequency levels of the scoring list and presents all words not found in the JACET 8000 list as "off-list types". This list of off-list words was then double-checked for proper nouns, wrong lemmatizations and variants of British or American spelling to eliminate any mistakes in the scoring procedure. Double-checking the results of the cleaning procedure in this way is one way to ensure the comparability and reliability of Lex30 results during scoring and especially important in this study because no further procedures for establishing reliability could be used, since a test-retest procedure was not practically feasible and the calculation of internal consistency measures such as Cronbach's α was deemed incompatible with the test format of word associations.[131]

Finally, the Lex30 score was calculated: all types occurring among the 1000 most frequent words in the JACET 8000 list were awarded 0 points. All other responses, thus, types in the 2K list and above and types not found in the JACET 8000 list, received 1 point.[132] In addition, the number of misspelled, but acceptable words was counted in the original Excel file. An adjusted score was calculated by deducting 0.5 points for each incorrect spelling from the JACET score. These two scores as well as further information such as the number of responses, types and tokens produced as well as the number of mistakes made were then summarized for each participant and form the basis of data analysis.

In the present study the students produced a mean number of 99.51 answers ([97.03, 101.91], *SD* = 17.06) in response to the 30 cue words on the Lex30 test. In the cleaned samples the mean number of tokens then was 91.49 ([88.85, 93.87], *SD* = 16.76) and the mean number of types was 80.17 ([77.88, 82.50], *SD* = 15.28). As mentioned above, the mean number of types is of special interest because response words are only counted once in the scoring process even if they were produced twice or more often in relation to different cue words. The frequency-based analysis results in a mean Lex30 score of 38.23 ([36.51, 39.92], *SD* = 11.41). However, on average the students made 3.34 mistakes ([2.98, 3.75], *SD* = 2.58) while filling in the Lex30 test; therefore, the adjusted score is slightly lower with an average of 36.56 ([34.88, 38.22], *SD* = 11.49). An interesting question in this respect is whether students with lower (unadjusted) Lex30 scores made more spelling mistakes than students with higher scores. A visual analysis and the very low correlation of τ = .02 ([-.10, .13], *p* = .776) indicate, however, that this is not the case. For this reason, results based on the regular Lex30 score unadjusted for spelling are presented Chapter 6, which also increases comparability with previous studies (e.g. Alejo González & Piquer Píriz 2016; Fitzpatrick 2012; Moreno Espinosa 2010; Walters 2012).

---

[131] Although Fitzpatrick and Clenton (2010) present a result for Cronbach's and Meara and Fitzpatrick (2000) calculated split-half reliability, it is not entirely clear to me how such statistical measures of internal consistency apply to the Lex30. The test does not rely on the notion of correct answers, but rather elicits a sample of words to be analysed further; thus, the application of such statistical procedures seems inconsistent with the test design.
[132] Types not occurring in the JACET list were double-checked with other sources such as the Oxford English Dictionary (www.oed.com/) or the Merriam-Webster dictionary (www.merriam-webster.com) to make sure they actually existed and were spelled correctly.

The results for productive vocabulary knowledge based on Lex30 are further described in section 6.4.4 before investigating possible connections to EE and other influencing factors. In addition, section 6.4.7 presents an additional analysis of the Lex30 data using vocabulary lists based on the participants' English coursebooks. This novel approach to analysing samples elicited with the help of Lex30 is used to explore characteristics of words likely to have been learned outside school.

5.3.6 Quantitative analysis of the combined dataset

The quality of quantitative analyses in L2 research has lately been criticized in a number of publications (e.g. Lazaraton 2000, Plonsky 2013, Plonsky 2014). They show that quantitative studies in the field strongly rely on null-hypothesis testing without considering the flaws of such procedures (Plonsky 2014: 461) and that statistical issues are aggravated by small sample sizes, low power and little reporting of standardized effect sizes (Plonsky 2013: 678).[133] In addition, frequently not even descriptive statistics are reported (Lazaraton 2000; Plonsky 2014). While word limits in academic publications most likely exacerbate these problems, it is alarming that conclusions from quantitative L2 research have to be drawn relatively frequently from insufficiently reported analyses.

Since the current study does not have the problem of a stringent word count, every attempt is made to follow the current state of the art for statistical analyses as set out, for instance, by Larson-Hall and Plonsky (2015). They recommend reporting descriptive statistics, effect sizes and reliability measures, analysing and presenting visual displays of data, sharing primary data where possible, and making an effort to interpret findings synthetically from a meta-analytic perspective. As a consequence, the quantitative analyses reported in Chapter 6 include descriptive statistics, effect sizes and confidence intervals, and methodological aspects of the study such as reliability are investigated in several ways. Visualizations of the data were inspected before applying any statistical procedures and are frequently presented to illustrate results. As far as data sharing is concerned, legal restrictions by the educational board of Vienna apply, which make the publication of the primary data difficult. Due to the exploratory nature of this study overall, the statistical analyses are also exploratory and investigate several factors with regard to their influence on EE and vocabulary size. Before setting out the results of the quantitative analyses in the following chapter, an account of the procedures is given below.

In general, statistical analyses were carried out using the software packages SPSS (Version 22.0, IBM Corp. 2016), R (Version 3.5.1, R Development Core Team 2018) and RStudio (Version 1.2.1335, RStudio Team 2018).[134] In this study, SPSS was mostly used for data entry and variable transformations (see also section 5.3.5), while the majority of statistical analyses were done in R

---

[133] Criticism against the dogmatic use of null-hypothesis testing has been voiced as early as 1957 (Larson-Hall 2016: 129), but has greatly increased in the last two decades (Cumming 2012, 2014; Plonsky 2015). One solution to the problems of null-hypothesis testing is to use "new statistics" (Cumming 2012, 2014), which is advocated by prominent figures in the field of quantitative L2 research (Larson-Hall 2016; Plonsky 2015; Porte 2010). This approach adds to null-hypothesis testing by using a-priori power analyses and by introducing more informative measures such as effect sizes and confidence intervals.

[134] A full list of references for all software programmes used is provided in Appendix A.

and RStudio.[135] Distributions and statistical assumptions were routinely checked before the application of any inferential test through data visualizations and descriptive statistics. Confidence intervals for descriptive estimates, such as the mean and median, were calculated using bootstrapping, an approach that resamples the data in simulation analyses.[136] 95% confidence intervals are reported corresponding to the customary alpha level α = 0.05; following the suggestions by Cumming (2012), they are presented in square brackets after the estimate for which they were calculated like in the following example: M = 40.17, [37.00, 43.95]. More detailed information on the procedures followed and on how the principles of the new statistical approach were put into practice is given together with the descriptions of the statistical analyses below.

Many of the analyses reported in Chapter 6 explore bivariate relationships and thus use correlations. Owing to the ordinal nature of a large part of the data and the presence of ties in many cases, it was decided to use Kendall's Tau (τ) as a correlation coefficient rather than the more commonly used Spearman rank order correlation (r_S).[137] Unlike r_S, which is actually the Pearson product-moment correlation coefficient applied to the ranks of the data, τ uses a less restrictive notion of correlation. While r_S requires a trend to be linear, τ is based on the assumption that if "the correlation between variables X and Y is strong if on average, there is a high probability that an increase in X will be accompanied by an increase in Y (or decrease in Y)" (Puka 2011: 713) and thus does not require linearity. In addition, τ is less sensitive to a large number of ties (Field, Miles & Field 2012: 225), especially in the version of Kendall's Tau-b, which accounts for the number of ties (Puka 2011: 714). However, disadvantages of Kendall's Tau are that the correlation coefficient is generally more conservative and thus smaller than r_S for the same data and that the squared coefficient does not adequately describe the proportion of shared variance (Walker 2003: 526), as is the case for R^2 calculated from r or r_S. Both these issues have to be borne in mind when interpreting bivariate correlations as effect sizes in the next chapter or when comparing the results of the present study to outcomes of similar research projects. Since Kendall's Tau can be interpreted as an effect size itself, no further effect sizes were calculated, confidence intervals for τ were constructed using the bootstrapping option of function kendall.ci() in package NSM3 (Schneider, Chicken & Becvarik 2018).

In addition to Kendall's Tau τ, the Spearman rank order correlation r_S is used for the calculation of correlations between variables which were measured on an interval scale data, but which do not show a normal distribution. The Pearson product-moment correlation coefficient r is only

---

[135] While SPSS may be more user-friendly at first glance, R is a more versatile tool which allows for relatively easy implementation of advanced statistical procedures such as bootstrapping and mixed-effect modelling, which are increasingly used in (applied) linguistics (Mizumoto & Plonsky 2016).

[136] More specifically, I used the boot.ci() function in package *boot* (Canty & Ripley 2017) with method BCA or perc after resampling from the data 2000 times. The bias-corrected and accelerated bootstrap (BCA) was the preferred method, but the percentile method was used for constructing confidence intervals for the median, if R reported the BCA to be unstable.

[137] As mentioned in section 5.3.5.1, in the present study rating scales are regarded as ordinal data and thus median values were used for the creation of summary variables. This procedure means the values can only vary between the five values of the original rating scale, which results in a relatively large number of tied ranks in bivariate correlations.

used as an effect size, it is preferred to other options because of its simple interpretation (Field, Miles & Field 2012: 57–58; Kirby & Gerlanc 2013). The interpretation of effect sizes follows the recommendations by Plonsky and Oswald (2014), which are based on a large meta-analysis of effects in quantitative L2 research. For the correlation coefficient *r* they suggest that .25 should be seen as small, .40 as medium and .60 as large in terms of effect. Furthermore, a special type of correlation also employed in the analyses is partial correlation, which "is the correlation between two variables with one or more variables partialled out of *both X* and *Y*" (Howell 2013: 528, emphasis in original). Partial correlations are thus useful for looking at the relationship between two variables while controlling for one or more mediating variables. Controlling the influence of a mediating variable for only one of the two variables results in a semipartial correlation (Field, Miles & Field 2012: 237; Howell 2013: 528). The squared semipartial correlation (sr^2) is used as an effect size in multiple regression (Larson-Hall 2016: 248; Tabachnick & Fidell 2013: 144).

In addition to correlations, differences between groups are also of interest in the analysis. For tests of difference between two independent sub-samples non-parametric Wilcoxon rank-sum tests were mainly used instead of independent-samples t-tests because the data often did not show a normal distribution.[138] For the Wilcoxon rank-sum test the effect size *r* can be approximated by calculating a *z*-score based on the *p*-value and adjusting it for sample size (Rosenthal 1991: 19 as cited in Field, Miles & Field 2012: 665). In the few instances in which more than two independent samples are compared, the Kruskal-Wallis test, the non-parametric equivalent of ANOVA, was used (Field, Miles & Field 2012: 674).

Far fewer comparisons were made based on paired samples; only for the comparison of the two different scoring formulae for V_YesNo two sets of scores drawn from the same group of participants are used. In this case, the Wilcoxon signed-rank test, a non-parametric version of the paired-samples t-test, was used. Due to the similarity in name this test is easily confused with the Wilcoxon rank-sum test for independent samples mentioned above, but there are important differences. While both tests work with ranks of the data, the rank-sum test ranks all data and then compares the sum of ranks in two conditions, as indicated by the name, whereas the signed-rank test uses the ranks of the differences between two related sets of scores. An effect size *r* can however be calculated in the same way for the signed-rank test as for the rank sum test (Field, Miles & Field 2012: 667–673).

Multivariate analysis was conducted to explore the combined effects of predictor variables on the receptive and productive vocabulary size test scores. Originally, the use of linear mixed models was planned to account for the hierarchical structure of the dataset with classes nested in school and schools nested in classes; however, the number of students in some groups was so low that the estimation of mixed effects was not possible: when fitting a mixed effects model to

---

[138] It is worth pointing out that this test is also known as Mann-Whitney U test, for example in its implementation in SPSS, because the Wilcoxon rank-sum test and the Mann-Whitney U test are actually equivalent (Field, Miles & Field 2012: 655).

the data computational issues occurred with regard to the random effects. A closer inspection of the model building procedure showed that even if only two fixed effects and random effects at school level were specified problems of singularity emerged during estimation of random effects. Since a valid estimation of both fixed and random effects was not possible with the collected data, standard multiple linear regression models were used. These were constructed separately for the V_YesNo scores and Lex30 scores as response variables, the results are presented in section 6.4.3 and 6.4.6 respectively.[139]

Since multiple regression already represents a relatively sophisticated multivariate statistical technique, sample size is an issue that needs to be discussed. Diverse recommendations can be found in the literature ranging from frequently reported rule of thumb of ten to fifteen participants per predictor variable (Howell 2013: 527; Stevens 2009: 117–120) to at least 30 participants per predictor variable (Porte 2010: 171) and many variations in between (see also Field, Miles & Field 2012: 273–274, Larson-Hall 2016: 238–240). For the present study these different guidelines would have resulted in recommended sample sizes between 60 and 180 participants. Cohen et al. (2003: 92) provide more persuasive guidelines based on power analysis for multiple regression. According to these, 123 participants are needed to achieve a power level of 0.80 for an expected small effect of $R^2 = .10$ and six predictor variables (Cohen et al. 2003: 651; Larson-Hall 2016: 239–240).[140] The sample size of the present study thus is sufficient for the application of multiple regression, but see section 6.4.3 for further discussion based on the actual results. In addition to needing a large enough sample, multiple regression also has a number of further assumptions that need to be fulfilled (Larson-Hall 2016; Tabachnick & Fidell 2013). Details on the examination of these assumptions and their interpretation are given directly with the presentation of the first model in section 6.4.3.

5.3.7 Summary

Following the introduction of the MMR research design in section 5.2.2, this section provided in-depth information about the quantitative strand of the study. Careful considerations guided the sampling procedure for the quantitative strand since it also constitutes the basis for the smaller qualitative strand in line with the overall sequential QUAN-qual design. Taking into account practical limitations in terms of geographical scope, comparability with other studies, and the results of the pre-pilot studies it was decided to focus on Viennese students attending grade 10 of public academic secondary schools that do not provide special forms of English teaching such as CLIL. The schools eligible for inclusion in the study were then categorized according to the percentage of non-German speaking students and teachers at schools from each category were contacted about taking part in the study. Decision to take part lay with teachers and heads of

---

[139] Note that standard multiple regression models using the lm() function in R produce exactly the same estimates for fixed effects as the mixed effects models using lmer() from package *lme4* (Bates et al. 2015) for which random effects converged to singularity.

[140] Although the original plan was to use a linear mixed effects model, power analysis and considerations of sample size were approximated on the basis of the more standard procedure of multiple regression, as the calculation of power for linear mixed effects models is not straightforward and commonly involves computer simulations based on the specifics of a study (for further information see Westfall, Kenny & Judd 2014.)

school; hence, a convenience sample had to be used due to the practical limitations of school-based research. Overall, 239 students from 12 classes in seven schools were informed about the study and asked to take part; data was collected from 224 students participated after obtaining consent from both students and parents. After the application of exclusion criteria in relation to English as a home language, specialized forms of English education and stays abroad in English-speaking countries, data from a final sample of 201 participants can be used in the analysis.

Next, detailed information was provided on the construct and development of the Extramural English Questionnaire (EEQ) and the Extramural English Online Language Diary (EEOLD). Both are based on examples used in previous research, but especially the EEQ takes new aspects such as language awareness into account. Selection of the vocabulary measures proved to be a difficult undertaking, but in the end a decision was made to use V_YesNo as a measure of receptive vocabulary size and Lex30 for productive vocabulary size. Reasons for these decisions, in which the practicality of data collection in schools played an important role, are set out in detail in section 5.3.3.3. Quantitative data collection was mainly carried out using paper-and-pencil versions of the research instruments, only the EEOLD was administered online. The fact that students could not be reminded to fill in the language diary for reasons of anonymity is one reason why the response rate is considerably lower in the online data collection.

Preparing the questionnaire and diary data for analysis after data entry entailed different procedures for data checking, application of transformations, analysis of internal consistency for multi-item scales and the calculation of summary variables. Scoring procedures for the two vocabulary measures followed detailed protocols informed by the developers as well as other users of the two tests. In a final step, descriptions of all statistical analyses used to explore the dataset including the combined data from all four quantitative instruments were provided together with their rationale. The outcomes of these analyses are presented in Chapter 6.

5.4   The qualitative strand

This section presents further information on the qualitative strand of the mixed methods research design, which followed the quantitative strand in a sequential manner, as described in section 5.2.2. Although priority is given to the larger quantitative strand in this study, the qualitative strand adds valuable insights by allowing a deeper understanding of students' perceptions of extramural English and its relation to language, and more specifically, vocabulary development. In addition, the findings of the qualitative strand help to explain surprising results found in the quantitative part, showing that this QUAN → qual MMR design truly has explanatory benefits as pointed out by Creswell et al. (2003: 227).

In parallel to the description of the quantitative strand, the following sections provide detailed information on the sampling strategy and the recruitment of participants for the focus group interviews (5.4.1), the interview guide (5.4.2), the procedures for qualitative data collection (5.4.3), the participants in the final groups (5.4.4), transcription and preparation of the audio

data (5.4.5) and the qualitative content analysis (5.4.6). First, the research questions of this strand and the reasons for using focus group interviews are briefly laid out below.

To reiterate, the research question addressed in the qualitative strand is

> **RQ 5:** *What are Viennese upper secondary school students' perceptions of EE and its potential for language learning?*

This relatively open question was split into more specific sub-questions to guide the design and implementation of the qualitative strand:

> **RQ 5a:** *How do participants describe the importance of English in their everyday lives?*
>
> **RQ 5b:** *How do participants describe their own EE practices?*
>
> **RQ 5c:** *How do participants interpret and evaluate selected results of the quantitative part of the study?*
>
> **RQ 5d:** *How do participants describe learning from EE?*
>
> **RQ 5e:** *How do participants evaluate (the potential for) learning from EE?*
>
> **RQ 5f:** *How do participants describe their practices of vocabulary learning from EE?*
>
> **RQ 5g:** *How do participants describe the relationship between their out-of-school English activities and their English lessons at school?*

As suggested by these more detailed questions, the focus of the qualitative strand is to elicit both descriptions and evaluations of extramural English and its relation to (vocabulary) learning from Viennese adolescents. To this end, focus group interviews were conducted (see section 5.2.2) as

> [f]ocus groups provide a rich and detailed set of data about perceptions, thoughts, feelings, and impressions of group members in the members' own words. They represent a remarkably flexible research tool, in that they can be adapted to obtain information about almost any topic in a wide array of settings and from very different types of individuals (Stewart, Shamdasani & Rook 2009: 590).

Focus group interviews are frequently used to explore topics about which little is known and can also help to clarify and interpret the results of quantitative approaches: "In this use, the focus group facilitates interpretation of quantitative results and adds depth to the responses obtained in the more structured survey" (Stewart, Shamdasani & Rook 2009: 590). In addition, group interviews allow the collection of qualitative data from several participants in a relatively short amount of time compared to individual interviews (see Stewart & Shamdasani 2015: 45). Focus groups are also more appropriate than individual interviews for this study from a practical viewpoint: in the school setting group interviews were the most feasible option to gather the views of several participants outside class time and the prospect of not being interviewed on their own lowered the students' inhibitions about taking part.

Although data collected from focus groups may not be as detailed as those obtained from individual in-depth interviews, they provide participants' emic view on the topics discussed in an ecologically valid way as participants are allowed to "respond in their own words using their own categorizations and perceived associations" (Stewart, Shamdasani & Rook 2009: 593). The emergent, interactive nature of group interviews can also present difficulties, as responses are not independent of one another and group dynamics may influence the content of the interview

and even introduce bias (Stewart & Shamdasani 2015: 48). At the same time, interactive group interviews allow participants to build on each other's responses, which may create a "synergistic effect" (Stewart & Shamdasani 2015: 45) leading to the introduction of new arguments or the uncovering of differences in opinion between participants. In the present case, group dynamics could be of special importance as participants knew each other previously and therefore power relations unknown to the researcher may have been at work.

5.4.1 Sampling and recruitment of participants

Section 5.3.1 described the sampling strategy used for the quantitative strand, which also forms the basis for qualitative sampling, since the qualitative sample is nested within the larger quantitative sample. The population of the qualitative strand thus consists of the 201 students in the final quantitative sample. Following the sequential mixed methods sampling strategy (Teddlie & Yu 2007) described in section 5.2.2, the second qualitative sample was selected from these through purposive sampling based on the quantitative results. Due to practical limitations of time and resources, it was decided to limit the sample to students from one class per school so that a maximum of seven focus groups would be conducted.

A list of preferred candidates was established based on the results of the EEQ to include different participants and their viewpoints in the focus groups, similar to the strategy of extreme case sampling (Dörnyei 2007: 127–129). For the selection of the preferred candidates, quantitative data regarding the number of daily and weekly EE activities and students' assessment of where they use English most were used as well as gender and the number of home languages. However, because the focus groups had to be conducted outside class time and participation was voluntary in practice a convenience sample with volunteers from the chosen classes had to be used. In terms of sample size, I decided to ask five students in each class to participate in the interview to ensure that discussion among students could be generated even if one of the volunteers were absent on the day of data collection or did not obtain parental consent.

The recruitment procedure for participants in the focus groups followed the same steps in each class. During the first part of data collection for the quantitative strand, participants were asked whether they wanted to receive their results on the vocabulary measures (see section 5.3.4.1). Since all of the 12 classes expressed an interest, I visited each one again in the summer term 2017 and provided feedback on the results. This visit was also used to set up the focus group interviews in one class per school. If more than five students volunteered or if they were generally reluctant to take part, the prepared list of preferred participants was presented and the students who recognized their code were encouraged to take part, but no pressure of any kind was exerted and the fact that participation was voluntary was emphasized several times. In all but one school at least five students volunteered to take part in a focus group interview, so that all in all six groups were set up.

Following the selection of participants, the volunteers were briefly taken out of class to set a date for the focus group, clarify organizational aspects and provide them with information sheets for

themselves and their parents or legal guardians, which included a second informed consent form to ensure that the focus groups could be recorded on audiotapes (see Appendix A). In each group one student was chosen to act as a spokesperson who was asked to provide me with their contact details for further organizational aspects and to collect the signature slips confirming parental consent. After the focus groups were set up, the class teachers were contacted via e-mail and asked to reserve a room for the interview unless the students had volunteered to do so themselves. The spokesperson in each group was contacted a few days ahead of the interview to remind them about the time and location and to ask about the parental consent forms.

5.4.2 Qualitative instrument: the interview guide

Prior to preparing the interview guide for the focus groups, previous studies on extramural English or related topics that had used interviews were consulted (e.g. Anioł 2011, Grau 2009, Hyland 2004, Lai 2015, Lai, Zhu & Gong 2015, Mukundan, Khojasteh & Pearson 2009, Yap 1998), but for many of these the interview guides could not be obtained and for the remaining ones it was found that most of the questions were not adaptable to this particular research context or the set-up of the focus groups as part of an MMR project.

As a consequence, the topics covered in the interview guide were based on the specific research interests and questions addressed in the qualitative strand of the project and formulated by myself. In addition, they were informed by the descriptive results of the quantitative strand, which is why the interview guide was designed only after a preliminary analysis of the quantitative data (see section 5.2.2). As focus groups are a highly flexible research tool, they can take a variety of formats and more or less structured approaches (Morgan 2001: 147). The design used in this study is situated more towards the structured end of the continuum as the main aim is to receive answers to the key questions outlined in the interview guide, rather than exploring the topic in an open-ended group discussion (see also Stewart, Shamdasani & Rook 2009: 602).

The interview guide consists of five main topic areas sequenced from the general to the more specific:

- The significance of English in participants' everyday lives
- Extramural English activities and their own use of EE
- The significance of EE for learning English
- EE and vocabulary acquisition
- The relationship between EE and school

While it would have been interesting to include further aspects in the interview guide, such as participants' language learning histories and the development of their engagement with EE over time, constraints on interview time meant that the number of questions had to be limited (see also Krueger & Casey 2001: 7–8) and since this study is of a cross-sectional nature questions on participants' current use of and views on EE were prioritized.

Each of the five topic areas in the interview guide (see Appendix A) contains one or more key questions that are addressed in every focus group, while the remaining questions can be used

to probe for further details or gain insights on additional topics if time allows. In addition, visual inputs are provided in combination with three questions to gather feedback on the results of the survey and/or as a stimulus for discussion. Like the EEQ and the EEOLD, the interview guide was written in German because the focus groups should be conducted in a language that participants feel comfortable with and in which they can fully express their thoughts.

5.4.3 Qualitative data collection

Qualitative data collection took place within two weeks right at the end of the school year between 13 June and 27 June 2017 (see Table 5.10). The focus group interviews were all conducted at participants' schools, usually in an available classroom that had been reserved for this purpose. Since four of the six focus groups took place after school or during the students' lunch break, pizza or similar snacks and soft drinks were provided to make sure they had the possibility to eat before afternoon lessons and to reward them for taking part in the interview.

Interview	Date	Duration	Number of participants
FG SA02	13.06.2017	44:45	5
FG SF01	16.06.2017	50:25	3
FG SE01	19.06.2017	50:58	6
FG SD01	21.06.2017	46:04	5
FG SG02	26.06.2017	1:02:13	6
FG SC01	27.06.2017	1:17:26	5
Total		**05:31:51**	**30**

Table 5.10: Data collection schedule for the focus group interviews

At the beginning, tables and chairs were rearranged so that all participants could see each other and students were encouraged to relax and help themselves to the snacks and drinks provided. Care was taken to make the environment as welcoming as possible since "[t]he first few moments in a focus group discussion are critical" (Krueger & Casey 2001: 13) for the atmosphere and thus the success of the focus group. During initial small talk I set up the audio recorder, a Zoom H2n handy recorder, in the middle of the table and usually also a mobile phone to have a backup in case of a malfunction of the recording device. Once everything had been prepared, I welcomed the students again and thanked them for taking part before beginning the actual interview according to the prepared interview guide (see section 5.4.2). To help with the transcription a drawing of the seating arrangements was made during or right after the interview and a post-hoc reflection describing the atmosphere, the course of interview, any unforeseen events, especially prominent topics or any other aspects I had noticed was recorded as soon as possible after each interview as recommended by Krueger & Casey (2001: 16).

Although the interviews generally proceeded rather smoothly, a certain level of flexibility was always needed. Absent students limited the number of participants in the focus groups and unforeseen events, such as class trips or late arrivals, led to delays in the beginning or to rather abrupt ends. In addition, the time frame available as well as the time of day at which the interview was recorded seemed to influence students' communicativeness in its initial stages,

although participants generally were very forthcoming. Hence, despite some practical complications the focus group interviews provided detailed answers to my questions, entailed open discussions and yielded rich data on 15-/16-year-old Viennese students' perceptions of extramural English and related topics such as the role of English in Austria.

5.4.4 Participants in the focus group interviews

In total, 30 students from six classes took part in the focus group interviews (see Table 5.10). Anonymity was guaranteed to the participants and for this reason they were asked to choose a pseudonym, which was then substituted for their names in the transcripts. As no restrictions were imposed on the choice of pseudonym, some participants selected aliases which disguise their gender or ethnic background, such as famous people or animals. While pseudonyms that render key characteristics unrecognizable should be avoided if possible (see Kuckartz 2014: 171), in practice it was difficult to dissuade participants from their choices and their suggestions were therefore accepted to avoid antagonizing them. To aid the analysis and interpretation of the results, Table 5.11 displays each participant's chosen alias as well as key characteristics such as gender, the number of languages spoken at home and whether they use English more in lessons at school or in their leisure time according to the information provided in the EEQ. In addition to the self-chosen pseudonyms, any other names of people or places mentioned in the focus group interviews or identifying details such as the country of origin were also anonymized during transcription (see Kuckartz 2014: 171–172).

As can be seen from Table 5.11, each of the six focus groups included both male and female participants; in total 15 girls and 15 boys took part in the interviews. Concerning their language background 16 participants speak only one language at home, which is usually German, while 12 participants use two or more different languages with their families.[141] The majority of the participants use English more frequently outside school with 19 students indicating that they use it more in their leisure time and only nine reporting that they use it more during their lessons at school. This fact has to be kept in mind when interpreting the findings of the focus group interviews as it could have led to a tendency to evaluate EE rather favourably.

---

[141] Two of the 30 participants were absent when the Extramural English Questionnaire (EEQ) was administered; thus, no information on their language background and use of EE is available.

Focus group	Pseudonym	Gender	Home languages	Uses English more
SA02	Keanu (Reeves)	female	1	in lessons at school
	John W. (Watson)	male	1	in their leisure time
	Kira	female	1	in their leisure time
	Niall	male	1	in their leisure time
	Lia	female	1	in their leisure time
SC01	Maria	female	2	in lessons at school
	John M. (Miller)	female	3	in lessons at school
	Franz M. (Müller)	female	2	in their leisure time
	Jane	male	1	in their leisure time
	Paul	male	*not present at data collection session 1 (EEQ)*	
SD01	Emma	female	2	in lessons at school
	Vanessa	female	2	in lessons at school
	Mito	male	1	in their leisure time
	Susi	female	1	in their leisure time
	Anna	female	2	in their leisure time
SE01	Pinguin	male	2	in their leisure time
	John	male	2	in their leisure time
	George (Washington)[142]	male	1	in lessons at school
	Kirito	male	3	in their leisure time
	Franz	female	1	in lessons at school
	Karl	male	*not present at data collection session 1 (EEQ)*	
SF01	DJ	male	2	in lessons at school
	KingKong	male	2	in their leisure time
	Walküre	female	1	in their leisure time
SG02	Louise	female	1	in their leisure time
	Sebastian	male	1	in lessons at school
	Johannes	male	1	in their leisure time
	Elisa	female	1	in their leisure time
	Marie	female	2	in their leisure time
	Lukas	male	1	in their leisure time

Table 5.11: Participants in the six focus group interviews (pseudonyms)

5.4.5 Transcription and data preparation

In total, the recorded interview time amounted to 5 hours and 31 minutes for all six focus groups (see Table 5.10 in section 5.4.3). The recordings were then transcribed by a recent graduate of our department, who had experience with transcription of classroom and group discourse.[143] The transcription groups were transcribed verbatim; dialectal traces were translated into colloquial standard German and punctuation was adapted to fit the conventions of the written standard, but vocabulary and syntax were not changed (see also Kuckartz 2016: 166–170). Since no video recording is available, the statements were attributed to participants by their voices. The first question in each interview was answered by participants one after the other and I noted the exact order of the respondents. With the help of these speech samples participants could

---

[142] George presents a special case: it became clear during and after the focus group interview that he had only lived in Austria for a short time and that he had a very different approach to language learning from all other participants in the focus groups. He seemed under a lot of pressure to improve his English and reported being highly focused and intentionally engaging in activities. Since George does not represent a typical Austrian learner of English and his length of stay in Austria falls under the exclusion criteria presented in section 5.3.2, it was decided to remove him from the quantitative sample. Accordingly, his statements will not be focused on in the description of the qualitative findings either.

[143] An overview of the transcription conventions used can be found in Appendix A.

generally be identified in the audio recording, in the cases were the identity of the speaker remained unclear specific symbols (e.g. SX) were used.

During the transcription process the documents were formatted for subsequent importation into MAXQDA (MAXQDA Plus 2018, Version 18.0.2, VERBI Software 2017) as focus group transcripts (see Kuckartz 2016: 166). Once the interviews had been transcribed, I read and corrected the transcripts before importing them into MAXQDA. In total, the final transcripts of the six focus group interviews consisted of 4098 paragraphs, which practically correspond to speaker turns and 67,702 tokens, as shown in Table 5.12.

Interview	Number of participants	Duration	Number of paragraphs	Number of tokens[144]
FG SA02	5	44:45	464	9165
FG SC01	5	1:17:26	971	15652
FG SD01	5	46:04	559	10151
FG SE01	6	50:58	556	9601
FG SF01	3	50:25	730	10565
FG SG02	6	1:02:13	818	12568
Total	**30**	**05:31:51**	**4098**	**67702**

Table 5.12: Summary of numerical information on focus group interview data

5.4.6 Qualitative content analysis

A qualitative content analysis, or more exactly a thematic qualitative content analysis following Kuckartz (2014, 2016), was conducted using MAXQDA. As Schreier (2012: 1–9) points out qualitative content analysis is a way of describing the meaning of qualitative data with a focus on specific aspects that are determined by the research question(s). The specific type of thematic qualitative content analysis can be used with a range of interview data and is characterized by "a multi-stage process of categorizing and coding" (Kuckartz 2014: 69), in which both concept-driven and data-driven categories are combined in a coding system.[145] Therefore, several first-cycle coding strategies (see Saldaña 2016) were applied at different stages of the analysis.

Figure 5.11 displays the typical process of a thematic qualitative content analysis described by Kuckartz (2014, 2016), which was also followed in the present study. First, I carefully read through the data and noted any thoughts and comments in memos; in part, this was already begun while correcting the transcripts of the interviews, but a second cycle of initial work with the text was done once the data had been imported into MAXQDA. In a second step, the main categories were decided on: mainly these were pre-established by the topics covered in the interview guide, but two more categories were derived inductively from the interview data. In total, seven main categories were developed and applied to all of the six focus group interviews

---

[144] The information on types and tokens is taken from the word frequency tool in MAXQDA.
[145] In German this type of qualitative content analysis is termed „inhaltlich-strukturierende Inhaltsanalyse", which underlines the fact that during the analysis the content is structured thematically with the help of the category system.

in a first coding process, which mainly used the strategy of structural coding. This is a kind of "question-based coding" (Saldaña 2016: 98) and particularly suitable for data from semi-structured interviews and those involving several participants.
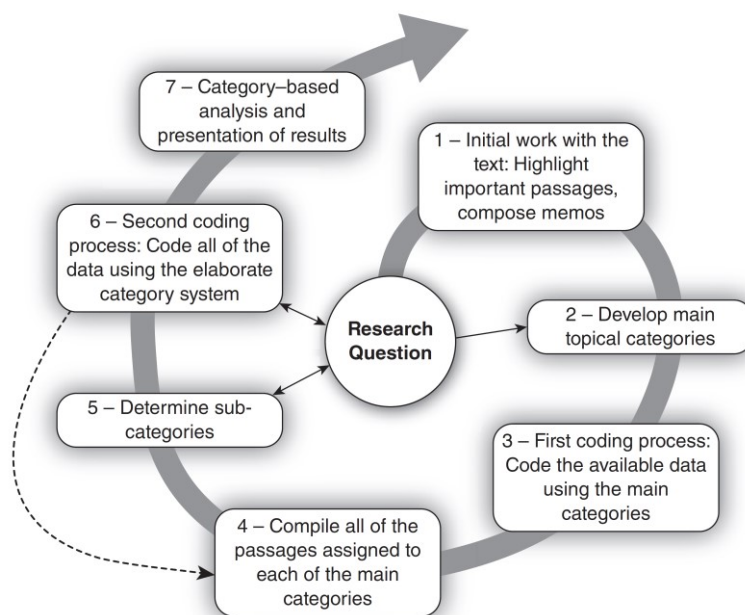
Figure 5.11: Thematic qualitative text analysis process taken from Kuckartz (2014: 70)

Following advice by Saldaña (2016: 79) in the first coding process the data were coded "in bulk": where possible, statements by single students were assigned to categories, but in all instances where several participants were discussing a particular topic large text segments were used as coding unit to preserve the context. Next, subcategories were generated for each of the main categories taking into account all segments that had been coded using a particular main category. Here, descriptive coding (see Saldaña 2016: 102–105) was applied to summarize the topics of participants' statements in short phrases,[146] and magnitude coding (see Saldaña 2016: 86–91) was used to classify participants' evaluations of learning experiences with EE and the learning potential of EE as positive, negative or mixed; therefore, the creation of subcategories was again both concept- and data-driven.[147]

The category system was then restructured before the final coding frame was documented in a codebook including names, descriptions and notes for all three levels of categories as well as examples from the data for the most detailed level of subcategories (see Appendix A). It should be noted that the subcategories could involve a second level of subcategories through application of subcoding (Saldaña 2016: 91–94) and hence the terms *main categories*, *categories* and *subcategories* are used to describe the three-tier structure of the coding frame.

---

[146] Unlike the general practice described in Saldaña (2016: 102–105) the application of descriptive coding in this study did not rely on single nouns, but short descriptive phrases to form a "categorized inventory" (Saldaña 2016: 104) of the data's contents.

[147] Because of the translation from German into English few original in-vivo codes, marked by quotation marks in the codebook, are included in the coding frame, but the wording of categories closely followed participants' choice of language in general to preserve their voices and reflect their language use.
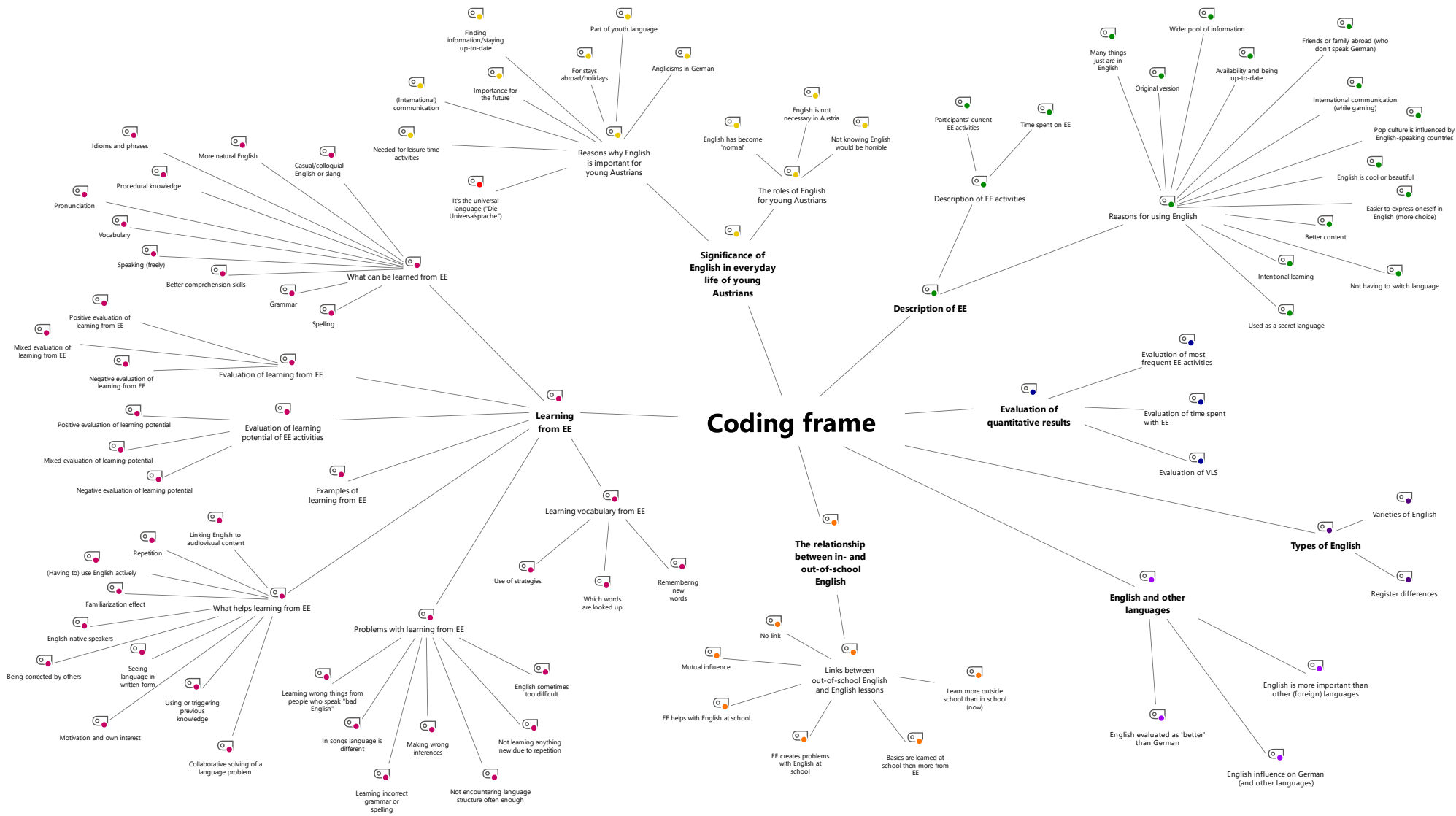
Figure 5.12: Visualization of the coding frame (using MAXMaps)

Figure 5.12 displays the finale coding frame, which consists of seven main categories, 20 categories and 68 subcategories. The main categories (formatted using small capitals) are mostly concept-driven and based on the main topics of the interview guide except for the two categories TYPES OF ENGLISH and ENGLISH AND OTHER LANGUAGES, which were derived inductively from the data. SIGNIFICANCE OF ENGLISH IN EVERYDAY LIFE OF YOUNG AUSTRIANS includes comments on the importance of English in participants' daily lives, DESCRIPTION OF EE features accounts of their EE practices and EVALUATION OF QUANTITATIVE RESULTS describes students' comments about selected findings of the quantitative strand in relation to EE practices and the use of vocabulary learning strategies. TYPES OF ENGLISH contains information on the varieties and registers of English students report encountering outside school and ENGLISH AND OTHER LANGUAGES reflects comments on the relationship between English and other languages in participants' environments. LEARNING FROM EE includes descriptions of several different aspects of learning from extramural English as well as evaluations of learning experiences involving EE activities and their learning potential. Finally, THE RELATIONSHIP BETWEEN IN- AND OUT-OF-SCHOOL ENGLISH features descriptions of current links between EE and English lessons at school.[148]

Each of the main categories encompasses two or more categories, which in turn may contain several subcategories, so that the complex coding frame depicted in Figure 5.12 presents a fine-grained analysis tool that allows exploring the data from different angles. The full codebook, which describes the coding frame including examples and notes, can be found in Appendix A. Once the final version of the codebook had been established, the complete dataset was checked again and coded using the full system of categories. Since the categories and subcategories had been generated by inspecting all segments included within a main category across interviews, in the second coding process the categories were checked and applied to one focus group at a time while listening to the audio recordings. In this final coding round interview segments could be assigned to several categories at once following the coding strategy of simultaneous coding (see Saldaña 2016: 94–97). Consequently, no claims concerning unidimensionality are made for the coding frame used in this study, as the categories and subcategories within one main category (also called 'dimension') are not mutually exclusive and data segments or units of coding were sometimes assigned to more than one subcategory within a given main category.[149]

After the second coding process had been completed, one focus group interview (SD01) was coded again five months later to establish the degree of consistency of the coding process. For practical reasons it was not possible to engage a second coder to establish intercoder agreement; therefore, intracoder consistency was calculated after the aforementioned time lapse. For this purpose, interview SD01 was coded again but rather than applying different levels of the coding

---

[148] This main category originally included a second category entitled "How in- and out-of-school English are and/or should be integrated by teachers". However, in the end the inclusion of an additional pedagogical focus in this thesis was deemed infeasible for reasons of scope; therefore, these data are not included in the analysis.

[149] Some scholars such as Schreier (2012, 2014) regard unidimensionality and mutual exclusiveness as requirements and key features of qualitative content analysis, whereas Kuckartz (2016: 43) takes a more liberal view and argues that in qualitative content analysis meaning units that contain enough context to be understandable when separated from the text should be chosen as unit of coding, and that therefore coded segments may well overlap or be nested within each other.

frame sequentially, main categories, categories and their subcategories were coded at the same time. A comparison of the original codings and the second round of coding shows that in terms of code occurrence there was a 95.1% overlap and in terms of code frequency 81.4% matched. Consistency at the segment level was much lower with 67.8% but this is largely due to the fact that I did not work with strictly segmented coding units as advocated, for instance, by Schreier (2012), so that the beginning and the end of the coded units differed in the first and second round of coding. Overall, however, the second coding process produced closely comparable results five months after the first coding was completed, which shows that the analysis of the focus group data is trustworthy.

In a final step, the focus group data were then summarized in thematic matrices using MAXQDA and analysed according to the main categories. In addition, relationships between different categories were explored verbally and through visual representations. No case-related summaries for individual participants were written because they are difficult to produce for interviews with multiple participants since audio recordings do not include information on nonverbal communication, such as gestures and facial expressions, which are frequently used to express agreement or a difference in opinion. As a consequence, the presentation of results in Chapter 7 is based on topics rather than on individual cases.

This section provided detailed information on the qualitative strand of the MMR study. Focus group interviews were chosen as the method of qualitative data collection because they give space to students' own voices, enable discussions among peers, and allow gathering data simultaneously from several participants. The participants in the focus groups were recruited from the larger quantitative sample and all volunteered to take part, although a wish list of ideal candidates based on the quantitative data was also used. All six focus groups use the same interview guide, which contains key questions on five topical areas as well as possible follow-up questions. Overall, 30 participants from six different schools took part in the interviews and provided insights into their perceptions of extramural English and its potential for (vocabulary) learning. The audio data were then transcribed and analysed using thematic qualitative content analysis, the results of which are presented in Chapter 7.

5.5   Summary

This chapter has introduced the empirical study which aims to map the landscape of extramural English in Vienna, Austria, to investigate the relationship between EE and vocabulary knowledge, and to explore learners' perceptions of EE and informal vocabulary learning. For this purpose, a cross-sectional mixed methods study was conducted with students in academic upper secondary schools. More specifically, this study uses a sequential QUAN-qual research design to confirm, complement and enhance the larger quantitative investigation with a more in-depth qualitative exploration. Mixing of quantitative and qualitative methods is present in the aims and rationale of the study, the sampling procedure used, the development of the qualitative instrument, the data analysis and the integrated discussion of results. Furthermore,

the detailed descriptions of the two strands show that the research design was carefully developed and implemented.

In the quantitative phase of the project data was collected from 224 Viennese students in grade 10 using four instruments. The Extramural English Questionnaire (EEQ) presents the core of the study and collects data on participants' use of EE, pertinent language variables, such as their home language(s), awareness and attitudes towards English, self-assessed English proficiency and vocabulary learning strategies, and sociodemographic variables in form of demographic data, SES, leisure activities and media access. The Extramural English Online Language Diary (EEOLD) complements the data collected on frequency of engagement with EE in the EEQ as it gathers more specific information on time spent with English outside school over the course of one week. In addition, V_YesNo (Meara 2015a; Meara & Miralpeix 2017) and Lex30 (Meara & Fitzpatrick 2000) are used to estimate participants' receptive and productive vocabulary size to be able to explore links between EE and vocabulary knowledge. Quantitative data were collected in 12 classes in seven academic schools in Vienna and low response rates for the EEOLD constituted the only problem encountered. After data entry, checking and transformations the data were analysed using descriptive and inferential statistical procedures following best practice guidelines in the field (Larson-Hall & Plonsky 2015; Plonsky 2015; Plonsky & Oswald 2014) as closely as possible.

The qualitative strand of the MMR design involved six focus group interviews with 30 students. This smaller sample is nested within the larger quantitative sample and descriptive quantitative results were taken into account for instrument development and sampling decisions. The interview guide consists of five main topic areas which explore the significance of English in learners' everyday lives in Austria, their engagement with and perceptions of EE, their evaluation of the potential of informal English use for language development in general and vocabulary acquisition more specifically, and their views on the use of English inside and outside school. The focus group interviews were generally carried out outside class time with students who had volunteered to take part. Following data transcription, they were analysed using thematic qualitative content analysis. Hence, both the quantitative and the qualitative data were first subjected to individual analyses, the results of which are presented in Chapters 6 and 7 respectively. However, in the spirit of mixed methods research and following the aims for using both quantitative and qualitative methods in this study, findings were then integrated in a second stage of analysis to compare inferences based on either strand and to develop meta-inferences drawing on both quantitative and qualitative data. These integrated results are discussed in Chapter 8 in light of previous research.

6 Results of the quantitative strand

This chapter presents the quantitative results on engagement with extramural English and its relation to vocabulary size. First, the sample of learners participating in the present study is characterized more closely (section 6.1) before providing a quantitative perspective on their perceptions of English (section 6.2). Next, data on the types and amount of contact with extramural English are presented (section 6.3). In the last part, the results on receptive and productive vocabulary size and its relation to extramural English as well as several other potential influencing factors are described (section 6.4).

6.1 Background information on participants

As described in section 5.3.5.1, a total number of 189 participants completed the extramural English Questionnaire (EEQ), which provides valuable information on students' background in addition to gathering data on their out-of-school English exposure. Thus, before turning to the results concerning extramural English the sample of learners participating in this study will be characterized more closely. First, information on demographic data and language background is provided (section 6.1.1), next their socioeconomic background and access to different media is described (section 6.1.2) and finally, participants' preferences in terms of leisure activities are set out (section 6.1.3).

6.1.1 Demographic information and language background

This section provides basic demographic information on participants and explores their different language (learning) backgrounds. In order to contextualize the background data on participants and to give an idea of generalizations possible from the sample, the information provided in this section will also be compared to larger-scale studies conducted in the Austrian context.

Characteristic		*N*	*%*
Gender	female	109	57.98
	male	79	42.02
Age	15	101	54.59
	16	66	35.68
	17	17	9.19
	18	1	0.54
Born in Austria	yes	159	85.03
	no	28	14.97
Number of languages spoken at home	1	84	45.16
	2	87	46.77
	3	11	5.91
	4	4	2.15
German as a home language	German used exclusively	75	40.32
	German used mainly	19	10.22
	German used partly	82	44.09
	German not used	10	5.38

Table 6.1: Participants characteristics

As can be seen from Table 6.1, more girls than boys took part in the present study with almost 58% of all participants being female. This ratio is however quite representative of the gender distribution in 10th grade in Viennese AHS, as in 2017 56.8% of a total of 5,178 students were female. Participants' age ranged from 15 to 18 with a mean age of 15.56 years, as is to be expected at this grade level.

The vast majority of participating students were born in Austria, but many have diverse backgrounds as indicated by the multitude of home languages present in the sample. The most frequently spoken home language is unsurprisingly German (60.5%), followed by Serbian (6.9%), Turkish (5.5%), Arabic (3.6%), Albanian (3.3%) and Bosnian (2.2%).[150] Yet, many more languages were named by students as a part of their linguistic environments at home: Armenian, Azerbaijani, Bambara, Bengali, Bulgarian, Chechen, Chinese, Croatian, Dari, Farsi, Finnish, French, Hindi, Hungarian, Italian, Korean, Kurdish, Luxembourgish, Macedonian, Malayalam, Pashto, Polish, Romanian, Russian, Slovakian, Spanish. The total number of 33 languages participants reported speaking at home is a testament to the linguistic diversity of Vienna.

Looking at the extensive list of languages above, it is unsurprising that many of the participants come from multilingual families. As shown in Table 6.1, more than 50% of students stated that they speak more than one language at home and the majority use two languages with their family members. It is not necessarily the case that one of these home languages is German, despite being the most frequently spoken language in Austria. Therefore, data on languages spoken at home were further analysed with regard to German, which has great significance as the default language of schooling. Manual coding of the home languages listed by participants resulted in the figures given in Table 6.1, which indicate that for nearly 95% of participants German does play a role as a home language, but the degree appears to vary considerably according to the students' self-report data: just over 40% identified German as the exclusive home language, about 10% use German as the main language in combination with one other language and the majority of nearly 45% use German to varying degrees as one of several home languages.[151] In line with the multitude of home languages given and the number on participants who reported using more than one language in their family, these results again underline the high proportion of multilingual participants in this study.

In addition to home languages the participants also study languages at school. As is common in Austria and particularly in the AHS school type, all students study English as their first foreign language (see section 4.2.1). What is interesting here are the self-reported data on age of onset for learning English, which are presented in Table 6.2. Several participants (10.75%) claimed to have started learning English between the ages of two and five, which suggests that they

---

[150] These figures are based on EEQ item 6a, which asked for home languages in relation to people living with the participants. The percentages given are based on the first language spoken with mother and father.
[151] The labels for German as a home language are to be interpreted as follows: 'German used exclusively' means that German is the only home language mentioned by the student, if one other language is given, but German is spoken with all family members the language situation was categorized as 'German used mainly'; 'German used partly' resulted from descriptions of a combination of languages in which German was used with several or only one family member, and 'German not used' was chosen if German was not mentioned as a home language at all.

experienced some form of English teaching at pre-primary level, for instance in kindergarten. The bulk of students (60.75%) indicated that they began studying the language between the ages of six and nine, which correspond to primary school level in Austria. Almost 30% reported that they began studying English at the age of ten or later, which either means that these students did not live in Austria at the time and/or were taught at different foreign language at primary level, or indicates that these students felt that their 'actual' learning of English only commenced in secondary school (see section 4.2.1).

Age at start of English learning	frequency	%
2[152]	1	0.54
3	5	2.69
4	8	4.30
5	6	3.23
6	37	19.89
7	22	11.83
8	35	18.82
9	19	10.22
10	40	21.51
11	9	4.84
12	4	2.15

N = 186

Table 6.2: Self-reported age of onset of learning English
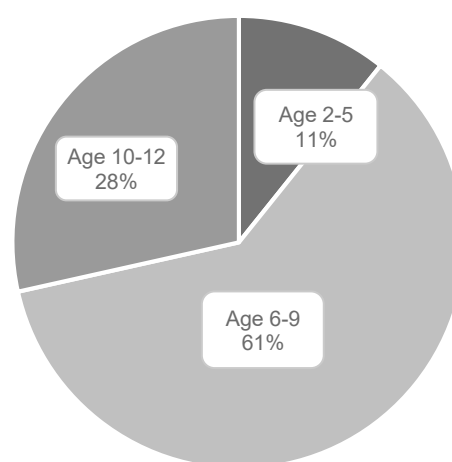
Figure 6.1: Self-reported age of onset collapsed into three age groups

In terms of their English proficiency, students in upper secondary AHS are supposed to have reached CEFR level B1 at the end of ninth grade (see sections 4.2.1 and 5.3.2). To get an impression of the language level among the participants in this study, a self-assessment scale based on the global CEFR descriptors for levels A2, B1 and B2 was used (see section 5.3.3.1). Participants rated their level of English for each of the four skills as summarized in Table 6.3.

Self-assessment according to skill	A2	%	B1	%	B2	%
listening	11	5.85	88	46.81	89	47.34
reading	19	10.11	51	27.13	118	62.77
speaking	32	17.02	101	53.72	55	29.26
writing	19	10.11	106	56.38	63	33.51

N = 188

Table 6.3: Participants' self-assessed English proficiency according to skill

Predictably, students seem to find the receptive skills of listening and reading easier than the productive skills. Over 60% chose level B2 for reading and a narrow majority also did so for listening, while more than 50% selected level B1 for speaking and writing. In general, Table 6.3

---

[152] There is no indication in other EEQ data as to why this participant began learning English at the age of two. German was given as the exclusive home language; thus, this participant does not have an English-language family background and was therefore not excluded from the sample.

suggests that most participants thought that they had already reached level B1 or even B2 at the time of data collection, a conclusion that is supported by further analysis of the data. In a second step the four ratings were aggregated in an index variable representing a sum score for each student (see Table B.1 in Appendix B for full results and details of the classification). These sum scores for overall language proficiency were then again categorized according to CEFR levels, which resulted in the following classification: 7 out of 188 students (3.72%) rated their overall English proficiency at level A2, 11 (5.85%) were in between levels A2 and B1, 80 (42.55%) judged their level as B1, 30 (15.96%) indicated they were between level B1 and B2, and 60 (31.91%) had reached level B2 according to their self-assessment. Hence, the vast majority of participants (90.37%) reported having attained or even surpassed the curricular target set for the successful completion of 10th grade in the course of the school year.

These data can be compared to the last school grades in English, which the students were asked to supply in the EEQ. A cross tabulation of students' last grades and self-assessed CEFR levels is displayed in Table 6.4. It indicates that self-evaluation does not contradict the grades for most students: the lower their grades were, the more students identified with descriptors for a lower CEFR level. Clearly, there are also exceptions, which is in line with the fact that neither school grades nor self-assessment scores are ideal indicators of actual language proficiency. Overall, however, the self-assessed level of English proficiency seems to provide a useful indication of learners' language level given the practical constraints of the study (see section 5.3.3.1).

	Level					
Grade[153]	**A2**	**A2/B1**	**B1**	**B1/B2**	**B2**	**Total**
1	0	0	4	7	20	31
2	0	2	18	6	19	45
3	3	6	33	13	13	68
4	4	3	24	4	7	42
5	0	0	1	0	0	1
Total	**7**	**11**	**80**	**30**	**59**	**187**

N = 187

Table 6.4: Cross-tabulation of participants' last grades and self-assessed CEFR levels (frequencies)

The majority of students also study languages other than English and German at school: 93.7% report learning a classical language such as Latin or a second foreign language. Table 6.5 displays how many of the participants study additional languages. As can easily be seen, Latin is still taught at many academic secondary schools in Austria, especially in those specializing in modern languages or the humanities. In the latter form, Ancient Greek is taught in addition to Latin, which is why eight participants also study this classical language. Even more common, however, is the teaching of modern second languages, which are represented exclusively by the Romance languages French, Italian and Spanish in this sample.[154] Hence, a typical participant in

---

[153] In the Austrian school system 1 is the best grade corresponding to 'excellent', while 5 is the fail grade corresponding to 'insufficient'.

[154] In addition to the main languages summarized in Table 6.5, Arabic, Farsi, Kurdish, Serbian, and sign language were mentioned as further languages studied by one student each. It appears that these languages were the

this study is likely to have studied Latin as well as a Romance language in addition to English, although the years of learning may differ considerably between languages.

Further languages studied at school	N	%
Latin	128	72.32%
Ancient Greek	8	4.52%
French	95	53.67%
Italian	22	12.43%
Spanish	28	15.82%

N = 177

Table 6.5: Languages studied at school in addition to German and English

Coming back to English, another aspect to discuss in terms of language background is participants' out-of-school exposure that is not included in the category of extramural English, as defined in section 2.2. Stays abroad and language camps, which would fall under the more general heading of language learning beyond the classroom, but are not the focus of EE, are examples of such exposure. In the present sample 13.8% have been to an English camp or an international summer camp where English was the main language at least once. In comparison, many more participants have been to English-speaking countries or used English while on holiday in non-English-speaking countries as shown in Table 6.6.

Characteristic		N	%
Number of holidays in English-speaking countries	none	84	44.68
	1 to 3 times	78	41.49
	4 to 6 times	18	9.57
	more than 6 times	8	4.26
Use of English on holiday in non-English-speaking countries	very little	27	14.44
	rather little	44	23.53
	rather a lot	77	41.18
	a lot	39	20.86

Table 6.6: Exposure to English during stays abroad

The figures in Table 6.6 indicate that more than half of the participating students have been to an English-speaking country at least once. While it is not clear how much English they used on these stays, it seems plausible that they had at least some language input and/or practice. At the same time, English, particularly in its function as a worldwide lingua franca, is often also used while travelling to countries where English has no official status. More than 60% of participants claim to have used English rather a lot or a lot on such holidays, which suggests that stays in non-English-speaking countries can also provide opportunities for language practice and development.

---

subject of extracurricular activities rather than regular language lessons because of the low number of references and the fact that they are not commonly taught at Austrian secondary schools (see section 4.2.1).

To sum up the results so far, we have seen that more teenage girls than boys took part in this study and that the majority of participants is 15 to 16 years old. More than half of the students come from multilingual homes in which a wide variety of languages is spoken. In addition, all participants study English at school and many also study Latin and a Romance language. The majority has been learning English for six to ten years and reports having a good command of the language on a self-assessment scale: 90% claim that they had attained CEFR level B1 at the time of data collection. In terms of language input that does not fall under the label of extramural English, slightly more than half of the participants reported visiting an English-speaking country at least once and around 60% have used English quite a lot on holidays in non-English-speaking countries. Finally, a small group of about 14% has also attended an English language summer camp at least once.

One last aspect to be described in this section is related more specifically to the research topic of this thesis, namely students' use of vocabulary learning strategies and their attention to new words in their surroundings. To collect data on the latter point, participants were given a set of statements in the EEQ (item set 5a) relating to their awareness of and attention to new English lexical items during their leisure time and had to indicate in how far these apply to them personally.

	strongly disagree (%)	rather disagree (%)	rather agree (%)	strongly agree (%)
I pay attention to new English words if they are important for understanding the content.	3.72	8.51	51.06	36.7
I intentionally pay attention to words that I don't know in English media.	9.57	40.43	33.51	16.49
I don't care about new English words – the main thing is that I understand the content.	16.49	46.81	30.85	5.85
I notice new English words even though I don't intentionally pay attention to them.	6.91	20.21	54.79	18.09
In my free time I am not interested in new English words.	36.7	43.62	15.96	3.72

N = 188

Table 6.7: Participants' awareness of and attention to new English words during leisure time activities

The results displayed in Table 6.7 show that more than 80% of the participants report being at least a little interested in new words during their spare time because they disagreed with the last statement (bottom row). It seems that most participants pay attention to new English words if they see them as important for understanding the content, but at the same time only understanding the content is not enough either, as is indicated by more than 60% disagreeing with the third statement. did not agree with this statement highlighting the importance of following a plot rather than understanding a new word. Some students intentionally pay attention to unknown words in English media in general with 16.5% fully agreeing and 33.5% partially agreeing with this statement. However, many more participants report noticing new English words even though they do not pay conscious attention to them, in fact, over 70% agree that this preliminary step to incidental learning happens unconsciously for them.

	almost never (%)	rarely (%)	often (%)	very often (%)	N
I think about what kind of a word it is (verb, noun...).	44.15	35.11	14.89	5.85	188
I try to separate the word into parts that I might now.	19.68	28.72	38.3	13.3	188
I think about if I know a similar word in other languages that I know.	12.77	17.02	40.96	29.26	188
If it comes up in a film or series, I try to guess its meaning with the help of the images and the story line.	7.98	8.51	42.55	40.96	188
I fit comes up in a text, I try to guess its meaning from context.	2.66	7.98	43.62	45.74	188
I look it up in a dictionary (also online or on the phone).	9.14	15.59	27.96	47.31	186
I ask somebody (parents, siblings, friends, ...) what the word means.	29.26	30.85	27.13	12.77	188
I don't do anything.	59.68	27.96	8.06	4.3	186

Table 6.8: Frequency of use for vocabulary learning strategies to discover new meanings during leisure time activities

The question then is what students do when they come across and notice an unknown English word during their leisure time. Table 6.8 displays the frequencies of use for a selection of vocabulary learning strategies (see section 5.3.3.1); encouragingly, almost 60% report that they almost always use some kind of VLS to discover the meaning of unknown lexical items, particularly if a new word is important for following the content as we have seen above. The most popular strategy for both audiovisual media and print media is guessing from context; over 80% of the students use it often or very often. Looking up a word in a dictionary, particularly online or with a smartphone app, is also done often or very often by three quarters of participants. As for language-based strategies to discover an unknown word's meaning, many students report thinking of other languages they know, but most hardly ever use the part of speech or word parts to analyse a new lexical item. Finally, the social strategy of asking somebody for help with unknown words' meanings is not very popular either with the majority of students, a fact that also comes up in the focus group interviews and will be revisited in section 7.4.

6.1.2 Socioeconomic background and access to media

In addition to demographic and language-related variables, the sample of participants can also be characterized by the students' socioeconomic background (see section 5.3.3.1). A first component of a scale of socioeconomic status is the educational level of parents: Table 6.9 indicates that more than 95% of the students come from families where at least one parent has achieved a qualification higher than a leaving certificate from compulsory school. Moreover, 76.87% passed the Austrian A-level exam: it is the highest qualification for 19.65%, while in 57.22% of the families at least one parent holds an academic degree.

Parental education given as highest ISCED level[155]	frequency	%
not completed any specific level of education/training	0	0.00
ISCED level 2 (obligatory school)	7	4.05
ISCED level 3V (school for intermediate vocational education, vocational school for apprentices)	25	14.45
ISCED level 3G (academic upper secondary school with A-levels)	19	10.98
ISCED level 5V (schools for master-craftsmen)	8	4.62
ISCED level 5V (college for higher vocational education with A-levels)	15	8.67
ISCED level 6 (BA programme)	23	13.29
ISCED level 7 (MA programme or comparable)	58	33.53
ISCED level 8 (PhD)	18	10.40

N = 173

Table 6.9: Highest educational attainment of students' parents according to ISCED level

These data suggest that the educational attainment of participants' parents is far above the Austrian average, as data from Statistik Austria show that in 2015 only 18.0% of Austrians had obtained a tertiary degree and 67.6% had successfully completed secondary school (Huber-Bachmann 2017: 35). Data concerning highest parental education in the PISA 2015 study show a similar picture: 6% of parents of the participating 7007 students had finished compulsory schooling, 37% had attended a vocational middle school or had completed an apprenticeship, 23% held the Austrian A-levels as highest qualification and 34% had obtained a degree in tertiary education (Salchegger et al. 2016: 88). Again, this indicates that the proportion of parents who successfully completed study courses in tertiary education is high in this sample, which means that the levels of educational attainment found in the participants' families are higher than in the wider Austrian population. However, it may be the case that the education levels among parents of students attending academic secondary schools are generally higher than in the population as a whole, as large-scale studies (e.g. Suchań & Breit 2016) have shown that education is hereditary to a large extent in Austria.[156]

In addition to education, data on parental occupation were also collected transformed to ISEI scores (see section 5.3.5.1). The highest score per family was again taken as an indicator of family background. Results concerning employment show that the vast majority of students' parents are employed or self-employed; only 18 parents do not work as they manage the household or are on maternity leave, retired, or unemployed. Among participants' parents in employment ISEI scores ranged from a minimum of 16.36 to a maximum of 88.96 with a mean of 59.41 [55.85, 62.98] and a standard deviation of 23.34. In general, these figures show that there is substantial variation in the sample with regard to occupational prestige, a result that was also found in the Austrian PISA 2015 study, in which ISEI scores ranged from 11 to 89 points (Salchegger et al.

---

[155] ISCED level 4 is not present in the table as it refers to "post-secondary non-tertiary education" (UNESCO Institute for Statistics 2012), which only applies to small-scale educational programmes in Austria (Austrian Agency for International Cooperation in Education and Research (OeAD) 2014) that were not explicitly mentioned in EEQ item 7c.

[156] Unfortunately data on students' socioeconomic background per school type are not publicly available.

2016: 87) and the mean score was 53.4 (*SD* = 0.47) for students without migratory background and 42.5 (*SD* = 0.91) for those with migratory background (Suchań & Breit 2016: 126). A comparison of these figures indicates that the range represented in this study's sample is similar to that observed in the much larger PISA study, but similar to parental education the mean ISEI score of students' families lies above the Austrian average.

As a third aspect of SES, home possessions were used as an indication of material wealth (Brese & Mirazchiyski 2013: 41, see section 5.3.3.1). The fact that 70.96% of the students reported having their own room in a capital city where housing space is limited and relatively expensive is another indication that the majority comes from a rather privileged socioeconomic background. In addition, Table 6.10 shows that students have access to a relatively large number of books at home, a variable that also indicates the availability of educational resources. In the PISA 2015 study 12.08% of students (N = 7007) reported that they had up to ten books at home, 43.63% up to a hundred, 18.72% up to two hundred, 15.61% up to five hundred and 9.49% stated there were more than five hundred books in their homes (OECD 2016: Table 1.6.2b, Annex B3). The figures in Table 6.10 are substantially higher concerning the upper categories of more than two hundred books which include more than 35% compared to approximately 25% in PISA.

Number of books at home	frequency	%
0-10 books	13	6.91
11-50 books	39	20.74
50-100 books	32	17.02
101-200 books	33	17.55
201-500 books	47	25.00
more than 500 books	24	12.77

N = 188

Table 6.10: Number of books available at students' homes

In addition to books, a range of media equipment is available in participants' homes, which is of special interest here as technical devices and especially an internet connection facilitate access to extramural English. Table 6.11 lists the media devices included in the EEQ and the proportion of students who reported being able to use them at home. Interestingly, access to the internet (99.46%) has surpassed the availability of computers (98.93%) and TV sets (97.35%) in this sample (compare section 4.3). This result is slightly counterintuitive but can be explained by the fact that nowadays computers are not absolutely necessary for surfing the internet as tablets or phones can also be used. At the same time television viewing is in decline because of a move towards online television services such as Netflix, Amazon Prime Video or Sky Ticket. More than half of the participants also have access to a DVD-player, a gaming console, a radio or stereo, a tablet and an mp3-player at home, whereas only about a quarter of families own an e-book reader.

Household possessions	frequency	%	*N*
computer (including laptops)	185	98.93	187
DVD-player (or Blu-ray)	154	82.35	187
e-book reader	49	26.20	187
gaming console	140	74.87	187
mp3-player	106	56.68	187
radio/stereo	131	70.05	187
tablet/iPad	125	66.84	187
TV set	184	97.35	189
internet access	184	99.46	185

Table 6.11: Availability of media devices at home

Table 6.12 provides information on the quantity of computers, mobile phones and TV sets at students' homes. Strikingly, and very much symptomatic of 21st century Austria, there is not a single household without a mobile phone and over 95% of households own three or more mobile phones. Likewise, over 50% of the families have three or more computers, but the majority only possess one TV set.

Household possessions	none	%	1	%	2	%	3 or more	%	*N*
computers (including laptops)	2	1.07	28	14.97	58	31.02	99	52.94	187
mobile phones	0	0.00	2	1.06	7	3.70	180	95.24	189
TV sets	5	2.65	87	46.03	60	31.75	37	19.58	189

Table 6.12: Number of computers, mobile phones and TV sets in students' homes

In sum, the number of media devices available at students' homes ranged from a minimum of 6 to a maximum of 15 in the present sample with a mean value of 10.69 with a 95% CI of [10.36, 11.01] and a median of 11 [10; 11].[157] This shows that generally participants have access to a wide range of media equipment which again implies a comparatively high SES of their families in terms of material wealth. The high SES level is also reflected in the possessions available for students' personal and private use, as illustrated by Table 6.13. All but one student own a smartphone and over three quarters have a personal computer at their disposal. Interestingly, less than half possess an mp3-player, but this may be due to the fact that smartphones have become all-round devices which can easily function as music players, too.

Student-owned possessions	frequency	%	*N*
smartphone	188	99.47	189
computer or laptop	148	78.31	189
mp3-player	87	47.03	185

Table 6.13: Possessions for personal use of students

The results concerning household possessions are not unexpected as Austria is a highly developed country and media access is widely spread. However, comparing the results of the present study to the figures for the wider population given in section 4.3 indicates a higher

---

[157] The number of media devices does not include internet access.

saturation with media devices in this sample than in Austria as a whole. The figures for media possessions in this study are consistently higher, but especially the level of computer ownership (99% compared to 87%) and internet access (99% compared to 85%) is striking. In part, this may be due to an age effect since it is unlikely that there are elderly people, who tend to show less affinity towards new technologies, among participants' parents. However, as we have seen, the sample generally exhibits signs of high SES in terms of parents' educational attainment and occupational prestige as well as the number of books available at home; thus, household possessions are another aspect underlining the socioeconomic advantages that the majority of participants enjoy. What these results concerning media equipment at students' home also suggest is that potentially they can use a wide range of diverse media to engage with extramural English.

6.1.3 Leisure time activities

Before turning to participants' use of English in their leisure time, it seems opportune to give an overview of their general spare time activities to contextualize the data on English activities presented in section 6.3. One of the research aims of this project is to map the landscape of extramural English, but it can only do so within a particular sample of Austrian adolescents. Hence, if many students in this sample favour a particular free time pursuit, this trend will also influence the activities they do in English; conversely, someone who does not do an activity will not engage in it in English either.

Table 6.14 displays the results for participants' general leisure time pursuits (for a full table including frequency counts see Table B.2 in Appendix B). It can easily be ascertained that there are three activities which more than 70% of the adolescents in this sample engage in almost every day: listening to music, using social networking sites and watching video clips on the internet. In addition, meeting friends, doing sports, watching films and series online and watching TV are also very popular.[158] In contrast, the least frequent activities, which more than half of the participants do almost never or only a few times a year, are (in ascending order of frequency): listening to audiobooks, going to the theatre, going to concerts, going to the cinema, playing multiplayer online games, watching films and series on DVD, making music, reading and playing digital games on their own. While the items on this list are not totally unexpected, it is surprising to find that gaming is a relatively unpopular free time activity in this group.

---

[158] Further leisure time pursuits mentioned by participants in an additional open question include: cooking, dancing, drawing, going out, making film clips, playing board games, programming, sleeping, writing stories, family activities or visiting relatives, and youth clubs (e.g. scouts, theatre group). Moreover, four participants also mentioned studying as something that they frequently do in their free time although this evidently was not the focus of the question.

Leisure time activity[159]	almost never (%)	a few times a year (%)	a few times a month (%)	a few times a week (%)	almost daily (%)	N
listening to music	1.59	0	2.65	8.99	86.77	189
watching video clips on the internet	1.06	0.53	3.72	19.68	75	188
using social network sites (e.g. Facebook, Instagram,...)	5.32	0	2.13	11.17	81.38	188
meeting friends	2.14	2.67	10.7	43.85	40.64	187
doing sports	4.81	2.67	19.25	52.41	20.86	187
watching films or series on the internet	8.47	5.29	23.81	29.1	33.33	189
watching TV	10.64	9.04	18.62	36.7	25	188
reading newspapers or magazines (also online)	13.44	10.75	29.57	32.8	13.44	186
making music (e.g. singing or playing an instrument)	48.13	6.42	9.63	18.18	17.65	187
playing games on a phone or tablet	18.28	16.67	30.65	17.74	16.67	186
listening to the radio	35.64	6.91	24.47	15.96	17.02	188
playing games online with others (e.g. Multiplayer Online Games)	47.59	11.23	13.9	14.97	12.3	187
playing games on your own (on a computer, console or online)	35.98	13.76	23.28	12.7	14.29	189
reading books or e-books	25.4	27.51	22.22	15.34	9.52	189
watching films or series on DVD	28.04	29.63	25.4	14.29	2.65	189
going to the cinema	6.95	57.22	32.09	3.74	0	187
listening to audiobooks	86.17	6.91	4.26	2.13	0.53	188
going to concerts	53.19	34.57	10.11	2.13	0	188
going to the theatre	60.64	31.91	7.45	0	0	188

Table 6.14: Frequency of general leisure time activities among participants

Moreover, not all teenagers participate in these leisure time activities to the same extent; in fact, for some activities there is a pronounced gender difference. As shown by Wilcoxon rank sum tests, boys play digital games significantly more often than girls, both on their own (W = 1624.5, p < .001, r = -.55 [-.64,-.44]) as well as with others (W = 883.5, p < .001, r = -.72 [-.78, -.64]) and also on portable devices such as phones or tablets (W =2582, p < .001, r = -.34 [-.46, -.02]). In contrast, girls read books (W = 5638, p < .001, r = -.27 [-.40, -.13]) and use social networks significantly more frequently (W = 5239.5, p < .001, r = -.29 [-.41, -.15]) and they also attend concerts (W =5098.5, p = .011, r = -.19 [-.14, .14]) and make music themselves (W = 5137, p = .006, r = -.20[-.34, -.06]) more often.[160] Despite these gender differences, overall a shift towards online activities appears to be taking place in this age group: two of the three most common activities are done online and films and series are consumed more through internet services than via more traditional media such as DVDs.

---

[159] The table is ordered by the combined frequency of 'almost every day' and 'a few times a week' as this definition of most frequent activities will also be used in the following sections.
[160] For an overview of all tests concerning gender differences in leisure time activities, please see Table B.3 in Appendix B.

To explore the topic of online activities further, participants were also asked to provide information on their daily internet usage and their favourite websites. Data show that out of 187 students who answered this questionnaire item, 186 or 99.47% use the internet every day. Table 6.15 presents the time spent online per day as estimated by participants grouped into four categories. It shows that the majority spend between two and four hours online, nearly a quarter of participants is online for less than two hours and roughly the same number fall into the third category of four to six hours. 16 participants report using the internet for more than 6 hours a day, which is a sizeable chunk of their free time considering the time necessary for basic needs such as sleeping and the hours spent at school. The mean time spent online per day is 3.95 hours ([3.65, 4.26], SD = 2.09) or 3 hours and 57 minutes, which is close to the median of 4 hours [3, 4].

Time spent online per day	frequency	%
0-2 hours	40	22.86
2-4 hours	78	44.57
4-6 hours	41	23.43
6-12 hours	16	9.14

N = 175

Table 6.15: Grouped estimates of time spent online per day

Having examined how long participants spend surfing the internet, the question remains what they do online. Although it was not possible to gather exhaustive information on online activities in this project, some indications are given by the websites students visit most frequently. Beginning with the most frequently mentioned these are: YouTube (142), Instagram (86), Google (62), Facebook (41), WhatsApp (37), Wikipedia (35), Snapchat (34), Burning series (28), Twitter (17), Netflix (16), and Amazon (10).[161] What is immediately obvious here is that not all responses given by students refer to websites in the traditional sense, some are more typically related to apps such as Instagram, WhatsApp or Snapchat. This is an indication that borders between different formats of online content are blurring and that phrasing EEQ item 1g in terms of websites may not have been the best choice since much content is available both in form of websites and applications for mobile phones or tablets. This is particularly true for social media and messaging platforms such as Instagram, Facebook, WhatsApp, Snapchat and Twitter, which are prominent among the most popular platforms. The most frequently accessed platform is, however, YouTube, which relates to the fact that watching video clips is one of the most common free time activities. The all-pervasive search engine Google and the online encyclopedia Wikipedia were also frequently mentioned, as were websites for streaming films and series, both legal ones, such as Netflix or Prime Video, and others operating in a grey area of the law, such as Burning series.

Websites that were mentioned less than ten times on their own were classified into categories (for a full table of codes and frequencies see Table B.4 in Appendix B). These show that students

---

[161] The number of mentions is given in brackets, please note that each student had the opportunity to name five websites, the total number of responses to item 1g was 695.

also browse news sites (17, e.g. orf.at, reddit and various newspapers), listen to music online (13, e.g. Soundcloud, Spotify), look at portals specializing in funny memes and satire (12, e.g. 9Gag), or use online dictionaries and translators (12, e.g. Google Translate, LeoDict). Furthermore, websites that host online games, allow users to watch others play, or provide a forum to discuss particular games were highly popular with 19 mentions. Students also reported further websites for streaming films and series (14), which again highlights the popularity of watching audiovisual media online as one of the most common free time activities. Other websites mentioned cover a variety of topics such as anime and manga, cooking, fan communities, movies, reading, shopping or travelling to name but a few, which shows that the adolescents in this sample had a wide range of different interests and passions.

The sociodemographic data presented in this section indicate that the sample of learners taking part in this study is representative of Viennese 10th grade students in terms of gender and age, but that the SES of participants' families as measured by highest level of education, occupational prestige and home possessions is higher than in other Austrian studies. In terms of language background there is great linguistic diversity in the sample: more than half of the participants come from multilingual families, but in the vast majority of these German is also used as a home language. Almost 90% of participants first began studying English at school, as is typical in Austrian, and over 90% judged their current English proficiency as level B1 or higher on a self-report scale, thus indicating that in their opinion they have achieved or even surpassed the curricular goal for English in 10th grade. In addition to English, a typical participant is likely to have studied Latin and a Romance language at school as well. Regarding contact with English through travelling, about 55% have visited an English-speaking country at least once and over 60% stated that they have used English a lot or rather a lot during stays in non-English-speaking countries. In contrast, only a small group of participants has ever attended an English language holiday camp.

In terms of access to extramural English the data show that students have a wide range of media devices at their disposal at home and their general leisure time preferences indicate that they make ample use of these, in particular to access online environments and watch audiovisual media. There are significant gender differences with girls reading more frequently and making more use of social media and boys engaging more frequently with digital games; overall however, gaming is not a popular leisure activity in this sample. All participants spend a lot a of their spare time online, in particular on the video platform YouTube, which clearly emerges as the favourite online platform. Findings further show that when participants encounter new English words in online and offline contexts they do pay attention to these, particularly if they are important for understanding the content, and that guessing from context, comparisons with other languages and using (online) dictionaries are the most frequently used strategies to discover the meaning of such new words.

6.2   Perceptions of English

Studying extramural English presupposes that participants come in contact with English outside their classrooms. However, to date little is known about where Austrian learners encounter English in their everyday surroundings (see section 4.4). It is therefore essential to this project to discover more about participants' awareness of English in their environment and their perception of its increased presence. EEQ item 1i asked participants to list the top three places in which they encounter English: "Please think about your everyday life in Austria: Where do you notice English most?" In the wording care was taken not to allude only to physical spaces and to include all aspects of students' daily lives. The total number of 521 responses was coded using an emergent category system, similar to the data on websites in the previous sections (for a full table of codes and frequencies see Appendix B Table B.5). Figure 6.2 visualizes the results for the eight most frequently named categories.
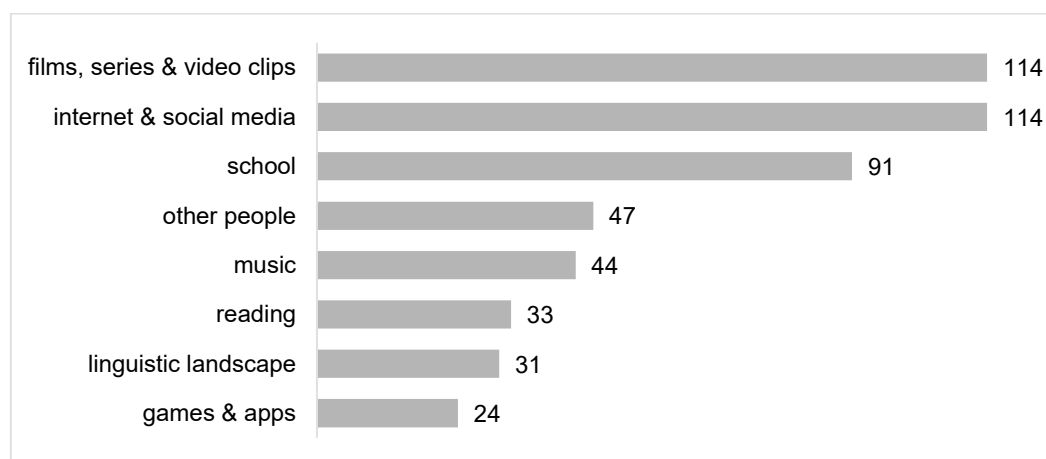
Figure 6.2: Where participants encounter English most often in their daily lives (frequency counts visualizing the eight most frequent categories based on data from open response items, N = 498)

As can clearly be seen, most responses fall into the two categories of 'films, series and video clips' and 'internet and social media' with 114 mentions each. The next most common answer was 'at school' with 91 mentions, which indicates the continuing importance of English teaching even though in students' perceptions it appears to have been overtaken by the internet as the most frequently noticed point of contact with English.[162] In addition, participants often come across English through communication with other people, music, reading, the linguistic landscape and games.

One more detailed finding regards the importance of video clips. Not only were YouTube videos mentioned 21 times as the most frequent site of encounter with English, but they also make up the largest part of the 'films, series and video clips' category as shown in Figure 6.3. Together with the findings that watching video clips is a highly frequent general free time activity and

---

[162] Of course, learners could potentially also encounter English at school outside their English lessons, for instance during breaks or in subject lessons. However, since none of the schools offered any CLIL programmes and several participants explicitly mentioned "lessons" or "English lessons" in their answers that are included in the category 'at school' it is very likely that encountering English at school mainly refers to encountering it during English lessons.

that YouTube is the website visited most often by students (see section 6.1.3), this result suggests that YouTube videos also have an exceptional status as point of access to English.
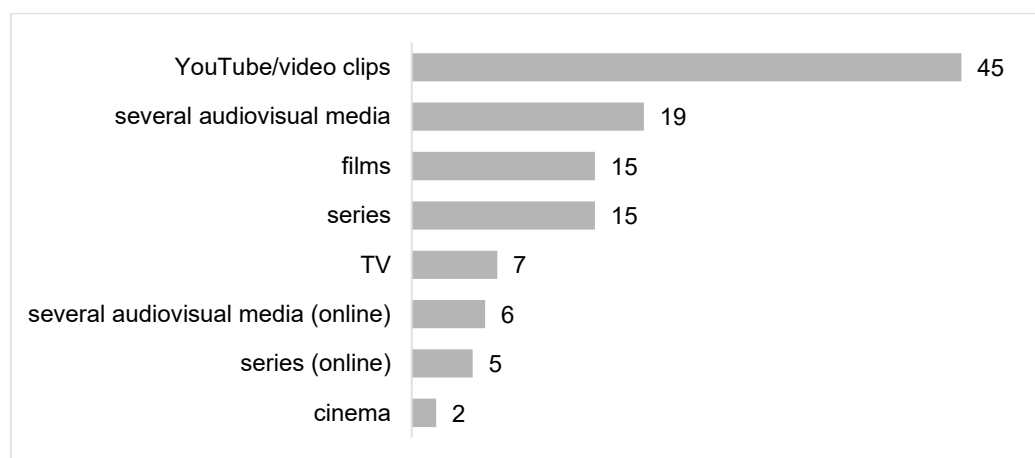
Figure 6.3: Where participants encounter English most often in their daily lives: constituent parts of the sub-category 'films, series and video clips' (frequency counts, N= 114)

Noticing English in the linguistic environment was also the subject of two closed items (4a and 4b). Data displayed in Table 6.16 indicate that the majority of participants hear or see English rarely in the places listed with the exception of public spaces, catering outlets and shops for which at least 50% of participants say they come across the language often or even very often. More pertinent to this research endeavour are the findings for item 4b, in which participants were asked whether they use English more in their spare time or in school lessons. Only 59 students (31.72%) stated that they use English more at school, whereas a much larger proportion of 127 students (68.29%) reported using English more for their leisure time pursuits. In light of the fact that in the past school was the most important opportunity for contact with English, this is a striking result.

Where do you see or hear English in everyday life?	almost never (%)	rarely (%)	often (%)	very often (%)	I am never there (%)	*N*
in the street	11.7	47.34	29.26	10.64	1.06	188
in public places (e.g. train station, parks, etc.)	6.38	38.3	40.96	13.83	0.53	188
in public transport	6.91	46.81	33.51	12.77	0	188
in banks, official offices etc.	15.59	46.77	20.97	4.84	11.83	186
in shops or shopping centres	9.04	40.43	32.98	17.02	0.53	188
in cafés, restaurants etc.	9.09	35.29	37.97	15.51	2.14	187
in places, where I go for my hobbies (music school, fitness centre)	21.81	39.36	16.49	12.77	9.57	188
at home	39.89	30.32	13.3	16.49	0	188

Table 6.16: Frequency of seeing or hearing English in everyday life (in %)

In addition to noticing examples in the linguistic environment, perception of English can also refer to evaluative stances towards the language. Therefore, an attempt was made to gauge participants' attitudes towards English or rather towards five specific functions of English: its

role in (their personal) future, in international communication, in young people's lives, in Austria and in relation to German. Each of these five constructs was transformed into four to five statements (see section 5.3.3.1), which were rated by participants on a four-point Likert scale (EEQ item 3). As reported in section 5.3.5.1, due to the low number of items per scale, none of the five multi-item scales reached adequate levels of internal consistency as measured by ordinal α; hence, the description below is based on agreement with the individual statements rather than the overall scales.

Table 6.17 presents the 21 statements which formed the five attitudes scales and participants' agreement with each individual item. As regards the first scale, over three quarters of participants agree that English is important for the future, with the first two statements receiving approval from over 90% of the students. Similarly, they recognize the importance of English as an international language as over 75% agree with three of the four statements. Significance for international travel appears to be particularly pertinent as agreement with the item "If you travel a lot, you must know English" was especially high.

Concerning the relevance of English for young people the vast majority of participants agreed that it was vital to learn English, but that it was also important for life outside school and overall more central to their lives than to those of older adults. The remaining two items elicited a more mixed response indicating that not every participant though that English was essential to having a good life as an adolescent, although over 50% rather disagree with the statement "Teenagers can get by well without English" and over 65% thought that it was more important for their spare time activities than for those of adults.

Looking at attitudes towards the role of English in Austria, a substantial majority of over 75% thought that it had a role to play and that English skills were not overrated. Participants were less decided if a good command of English was very important in Austria and they were almost split equally on the question of whether English was necessary for life in Austria nowadays. The same was true for most items regarding the relationship between English and German: while over 90% did not see English as a threat for German, opinions varied if it could be said to enrich the German language. Just over 50% thought English was more modern or sounded better than German, but the range of answers suggests that the aesthetic qualities of English are not the main incentive to engage with the language for many participants.

Item	Scale	Item text	strongly disagree (%)	rather disagree (%)	rather agree (%)	strongly agree (%)	N
AtE01_F1	Future	English is important for life after school.	1.59	2.12	19.05	77.25	189
AtE07_F2	Future	Knowing English will be very important for life in future.	0	3.7	29.63	66.67	189
AtE11_F3 [1]	Future	Knowing English is not necessary to have a good future.	32.28	49.21	13.23	5.29	189
AtE18_F4	Future	To get a good job later on, it is important to know English.	4.84	18.82	50.54	25.81	186
AtE02_I1	International	If you travel a lot, you must know English.	0.53	1.06	17.99	80.42	189
AtE04_I2 [1]	International	One can be international without knowing English.	18.72	57.75	19.25	4.28	187
AtE13_I3	International	With English one can make oneself understood wherever one goes.	0.53	4.28	49.2	45.99	187
AtE17_I4 [1]	International	One can communicate well with people from other countries without English.	13.83	54.26	23.4	8.51	188
AtE06_Y1	Youth	In the daily life of young people English is more important in their free time than it is for adults.	4.84	27.96	40.32	26.88	186
AtE10_Y2	Youth	For people aged 50 and over English is not as important as for young people.	5.35	14.44	40.64	39.57	187
AtE12_Y3	Youth	It is important for young people to learn English.	0	2.12	29.63	68.25	189
AtE14_Y4	Youth	English is important for life outside school.	2.15	7.53	42.47	47.85	186
AtE19_Y5 [1]	Youth	Teenagers can get by well without English.	9.68	45.16	38.71	6.45	186
AtE05_A1	Austria	English is also important within Austria.	2.12	21.69	48.15	28.04	189
AtE08_A2 [1]	Austria	English skills are overrated in Austria.	23.53	59.89	13.37	3.21	187
AtE16_A3	Austria	A good command of English is very important in Austria nowadays.	4.28	33.69	50.8	11.23	187
AtE21_A4 [1]	Austria	For life in Austria one doesn't need English.	13.9	37.43	37.97	10.7	187
AtE03_G1	German	English enriches the German language.	6.95	44.92	35.29	12.83	187
AtE09_G2	German	English is more modern than German.	9.73	25.41	38.92	25.95	185
AtE15_G3	German	English is a threat to the German language.	62.77	30.85	5.32	1.06	188
AtE20_G4	German	English sounds better than German.	13.37	20.86	26.74	39.04	187

[1] Negatively worded items were reverse coded for reliability analysis. In this table the original values are reported, for negative statements agreement with the overall construct of a scale is expressed by disagreement with the specific item.

Table 6.17: Proportion of agreement for each attitude item (in %)

In sum, participants in this study notice English most often on the internet and in social media as well as in films, series and video clips; with the latter appearing to be an especially frequent point of contact with English. School was also named as one of the top three places to encounter English by almost half the participants, but over 65% state that they use English more in their leisure time than in school lessons. In the linguistic landscape of Vienna, the places where about half of the students report coming across English often or even very often are public spaces, such as train stations or parks, as well as cafés, restaurants and shops. Regarding attitudes towards English, the students in this sample evaluate English positively overall: they agree that is important for the future and international contacts, especially travelling, and more importantly, over three quarters also believe that English is important within Austria, even if its relation to German is subject of a more nuanced interpretation.

6.3 Types and amount of contact with extramural English

This section will provide answers to the central question which kinds of extramural contacts Viennese 10th-grade students actually have with English during their leisure time. First, the results on the most and least popular EE activities are presented according to frequency (section 6.3.1). Next, findings on the amount of time spent with English based on the EEOLD are summarized (section 6.3.2) before turning to the reasons for engaging with English (section 6.3.3). Finally, differences in engagement with EE according to other background variables included in this study are explored (section 6.3.4).

Before presenting specific results on EE activities, it is interesting to relate English to other languages used during students' leisure time. Data from the EEQ suggest that overall participants use 43 different languages in their free time, but that – next to German – English has a special position among these. In EEQ item 2d participants were asked to estimate the proportion of time spent with each of their free time languages and descriptive analyses show that for 84.7% German is the language they use most in their free time, but for 50.8% English is in second place and for 33.9% it is in third place after a second home language. Although linguistic diversity is high, English thus appears to be the most commonly used language after German, while the remaining languages are used by individuals or small groups.

6.3.1 Extramural English activities according to frequency

Based on the 65 questionnaire items included in item set 2a of the EEQ (see section 5.3.3.1) participants report between 0 and 33 daily EE activities with an average of 9.43 ([8.54, 10.32], *SD* = 6.19). Only seven out of 189 students indicate that they do not use English for spare time activities on a daily basis; consequently, 96.3% of the participants engage in at least one EE activity (almost) every day. Looking at the larger unit of a week, participants' responses show a mean value of 16.87 EE activities ([15.79, 17.96], *SD* = 7.56), in which they engage at least weekly.[163] The figures range from 0 to 37 weekly activities, but only two out of 189 participants

---

[163] Weekly EE activities are defined as activities done at least a few times a week and therefore include the daily activities.

report not engaging in any weekly EE activities. The most common activities in which more than half of the participants engage at least weekly are presented in Figure 6.4. It contains both EE activities based on individual EEQ items and summary variables marked by the label 'overall' and bold type (see section 5.3.5.1).
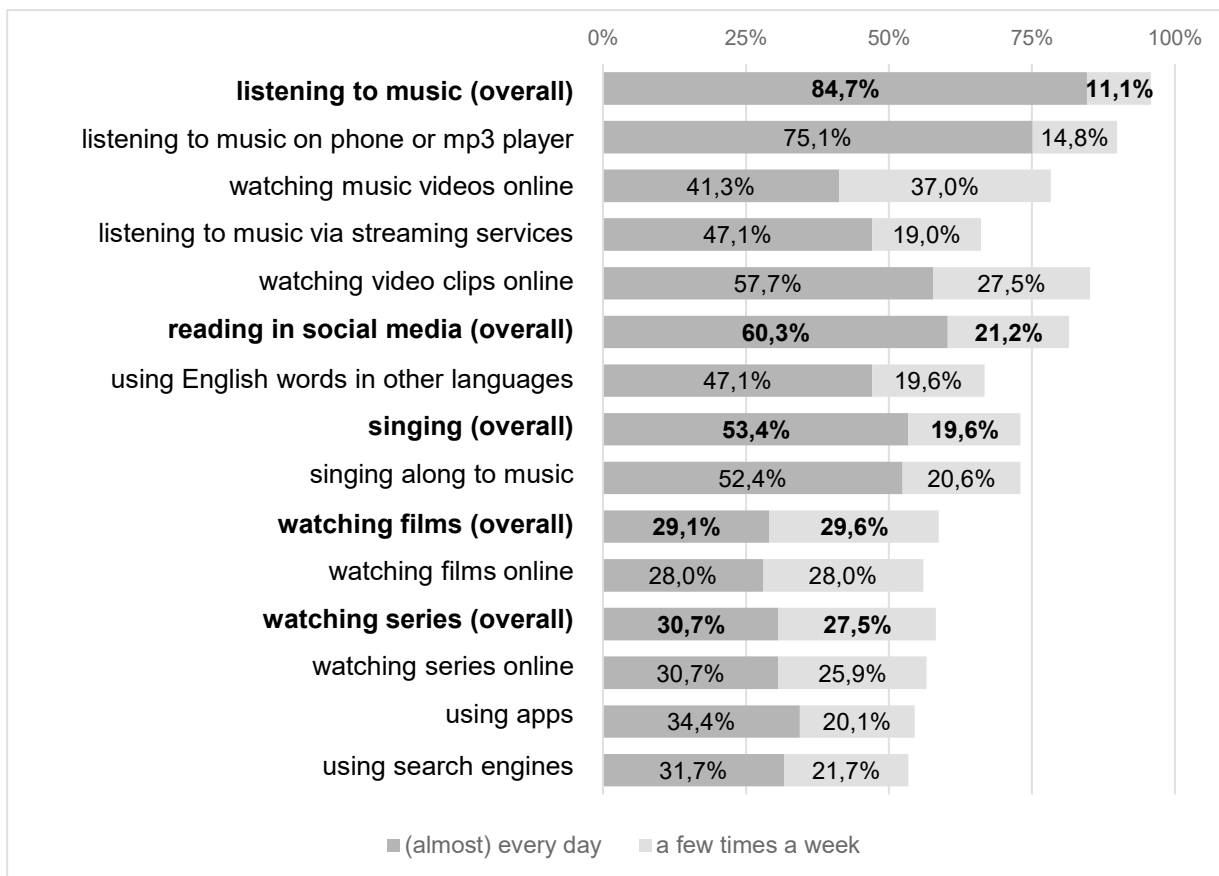
Figure 6.4: Most popular EE activities in which more than 50% of participants (N = 189) engage at least weekly (the label 'overall' and bold type indicate summary variables)

Listening to music is the most widespread EE activity with 95.8% of students reporting that they listen to English-language music at least a few times a week and nearly 85% almost every day. As can be seen from the next three bars in the chart, music is largely consumed via portable personal devices such as a phone or an mp3 player, but music videos as well as streaming services, such as Spotify or Soundcloud, are also frequently used. The second most popular EE activity is watching online video clips, which 85.2% do at least weekly and more than 50% on a daily basis. Next comes another online activity with reading in social media; here the individual items 'reading messages' and 'reading status updates or comments' did not reach the 50% threshold, but taken together they constitute the third most common EE activity.[164] It is worth noting that the top three EE activities correspond to the top three general leisure time activities – listening to music, watching video clips and using social media – in this sample (see section 6.1.3). This means that more than half of the students encounter English on a daily basis while pursuing their favourite spare time activities. Even if we disregard potential implications for

---

[164] Please see Table B.6 in Appendix B for full results on all EE activities.

language development for now, this is a fascinating result as it shows that for the majority in this sample, and by extension for many young Austrians, English is an integral part of their daily lives.

Next on the list of favourite EE activities is using English words in other languages, i.e. using English words or phrases while having a conversation in German, for example. This item was added to the EEQ due to students' suggestions during the pilot study (see section 5.2.3) and indeed 66.7% of participants in the main study agree that they use English words while speaking other languages at least a few times a week. Singing in English is also highly popular with 73.0%, but interestingly this figure can be explained almost exclusively by singing along to music (72.8%) rather than singing on one's own, which only 31.7% do on a weekly basis. Watching English-language films and series are also among the most frequent EE activities with 58.7% and 58.2% respectively, and as the two subsequent bars indicate most of this watching is done online. Finally, more than half of the students also use apps in English or search engines to locate information on the internet at least on a weekly basis.

What the most popular EE activities displayed in Figure 6.4 have in common is that they are predominantly done online and that most involve language receptively rather than productively. The results indicate that participants mainly encounter English through listening in their leisure time and that music, video clips, films and series are major sources of input. In comparison, the only activities that present the English language in its written form are reading in social media, using search engines and, to an extent, using apps. Search engines and perhaps certain apps also include some language production, but even if taken together with singing and using English while speaking other languages, the opportunities for productive language use in these common EE activities are limited. The same trend can also be observed when looking at EE activities in which participants engage at least a few times a month: Figure 6.5 displays activities which more than half of the participants report doing at least on a monthly basis, excluding those that were presented in Figure 6.4.

Figure 6.5 suggests that students are likely to encounter more written English in activities which most of them do a few times a week or a month. 78.4% read English lyrics at least a few times a month, 70.8% read articles and 68.2% read information texts, which were exemplified by instructions, recipes and Wikipedia articles in the EEQ. Following the trend towards online activities, most of this reading is done on the internet: the bars for 'reading articles online' and 'reading information texts online' are nearly as long as those for the summary variables which also contain the values for printed texts. Writing also plays a role in these weekly to monthly activities as writing English in social media, for instance in comments, status updates and messages, is done by 68.8%.
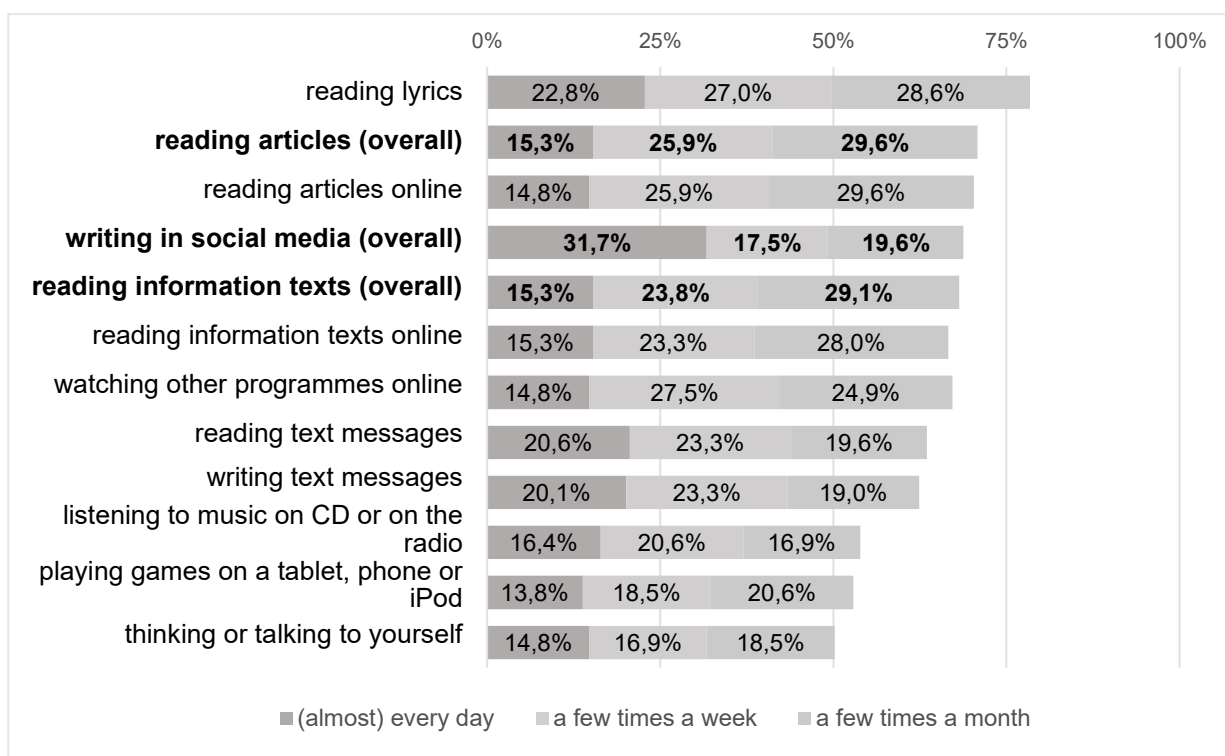
Figure 6.5: EE activities in which more than 50% of participants (N = 189) engage at least monthly (the label 'overall' and bold type indicate summary variables)

When examining the figures for daily and weekly activities more closely, it is interesting to see that apart from reading lyrics communicative activities show the highest frequency: 49.2% write English in social media on a weekly basis, 43.9% receive and read English text messages and 43.4% also write and send them a few times a week. This suggests that more than 40% of participants communicate in English with friends or other acquaintances relatively frequently. The remaining activities of watching programmes other than films and series on TV, listening to music on CD or the radio, playing games on a portable device and thinking or talking to oneself in English are not among the most popular, but are still undertaken a few times a month or more frequently by more than 50% of participants, as are the reading activities mentioned above.

Finally, the least popular EE activities are presented in Figure 6.6. Less than 20% of participants engage in these a few times a month and less than 10% a few times a week or more often. This may not be astonishing as many of the activities listed in Figure 6.6 are rather specialized; for instance, one would not expect adults, let alone teenagers, to watch English plays at the theatre more than a few times a year. Indeed, several of the least favourite EE activities correspond to generally infrequent leisure time activities as listed in Table 6.14: the majority of participants almost never go to concerts or the theatre, or listen to audiobooks. Consequently, it would be of rather little interest to look at how many learners do *not* engage in the EE activities listed in the figure above; in fact, it is much more enlightening to look at the small minority who do. For each EE activity there was at least one student in the sample who reported doing this activity at least a few times a month, which points to an impressive variety of English activities in this sample and makes it worth looking at this range of interests in more detail.

	0%	5%	10%	15%	20%

acting — 1,1%
writing comics/manga — 0,5% each
listening to audiobooks — 0,5% 2,1%
writing a diary — 0,5% 1,1% 2,1%
creating video clips — 1,1% 1,1% 1,6%
watching plays in the theatre — 4,8%
writing blogs/entries to forums — 1,1% 0,5% 4,2%
writing stories/fan fiction — 1,6% 2,6% 3,2%
watching films in cinema — 2,6% 5,3%
listening to programmes on the radio — 1,1% 1,6% 7,4%
listening to music at concerts — 3,2% 6,9%
reading e-books — 2,6% 4,8% 7,9%
writing e-mails — 4,2% 11,1%
reading articles in print — 1,1% 1,6% 13,2%
reading information texts in print — 0,5% 4,8% 11,6%
speaking via telephone — 1,6% 5,8% 10,0%
watching other programmes on TV — 1,1% 3,7% 14,3%

■ (almost) every day   ■ a few times a week   ■ a few times a month

Figure 6.6: Least frequent EE activities in which the majority of students engage only a few times a year or almost never (please note the change in scale in comparison to Figures 6.4 and 6.5)

Two participants, both female and male, report being involved in English-language acting a few times a month and when it comes to watching English-language plays at a theatre, nine participants, both male and female, claim to do so a few times a month.[165] One female participant likes listening to audiobooks almost daily, and another three females and one male do so a few times per month. Both boys and girls equally like to produce their own English-language video clips: two participants do so almost daily, two on a weekly and three on a monthly basis. In contrast, writing activities are favoured by female participants: three girls have a penchant for writing and drawing comics or manga in English with one girl claiming to do so almost every day, one a few times a week and one a few times a month. Seven girls write a diary in English: one almost every day, two a few times a week and four a few times a month. In contrast, English-language blogs and entries to forums are written by both female and male participants with two posting almost every day, one a few times per week and eight a few times a month. Stories and fan fiction writing is again more in female hands: three girls write stories almost every day, four girls and one boy a few times a week and three girls and boys each a few times a month.

Although this account has been limited to the eight least frequent activities and to those participants who engage in them at least a few times per month, it reminds us that behind the

---

[165] Whether these statements are true or not and whether students also thought of theatre performances they attended as part of school excursions is of course hard to say. However, since measures were taken to limit the chances of false reports or social biases as much as possible, there is a good chance that these results actually correspond to students' practices.

proportions and percentages presented above there are young people who pursue activities according to their individual preferences and that quite a few of them actually engage in seemingly 'unimportant' EE activities. This fact is also highlighted by examples students gave in response to the last item in EEQ item set 21 "I also do other activities in English, namely": one girl raps in English almost every day and one boy uses English for programming almost every day; one girl does sports, probably in an English-speaking team, a few times a week, another does research to prepare travels abroad several times a week, and finally yet another girl translates stories a few times a month. While interesting in themselves, several of these 'niche activities' involve much more language production in form of speaking or writing than the most common EE activities and they also require a certain level of language proficiency to be carried out successfully. This could make them especially interesting for language development, a point which will be taken up again in sections 6.4.2 and 6.4.5 and on the basis of the interview data in section 7.2.1.

Overall, however, this analysis of EE activities according to frequency shows that Viennese teenagers mostly encounter English in out-of-school contexts through listening and viewing. In addition, many of the most popular activities are carried out in online contexts, this is especially the case for reading: students report frequently engaging with English-language texts on social media, but they are unlikely to read print articles in newspapers or magazines. The results for writing activities show that slightly less than half of the students write status updates, comments or messages in social media at least a few times a week, but the majority does not write longer English texts in their spare time. Oral language production is mostly limited to singing, especially singing along to music, while interactive speaking is very infrequent in out-of-school contexts. Taken together, these results show that receptive language use is much more common in Viennese students' English leisure activities than productive language use and that online activities are a major source of EE input. At the same time it is also worth noting that with the exception of five items each EE activity was rated across the full range of the rating scale from 1 to 5. This implies that for all but five EE activities, at least one student engaged in it almost never, while another did it almost every day. Together with the niche activities presented above, this finding is testament to the highly individualized nature of participants' spare time activities and their use of English.

6.3.2 Time spent with extramural English activities

Data on the amount of time spent with English were collected with the help of the Extramural English Online Language Diary (EEOLD, see section 5.3.3.2). As discussed in Chapter 5, response rates for the online instrument were much lower than for the pen-and-paper instruments and the results presented below are based on 383 diary entries by 118 participants.

A first calculation of raw results shows that the mean time spent with EE across all diary entries was 231 minutes ([214.6, 248.2], *SD* = 167.4) or 3 hours and 51 minutes per day. The average time spent with EE can, however, also be calculated differently by first computing the mean time for

each of the 118 participants across all days for which they filled in the EEOLD and then taking the average across individuals. This second method takes variation according to participants into account and is therefore preferable; as a result, it will be used for all calculations of mean times in the following. According to the second method of calculation, mean time spent with EE activities per day was slightly higher with 248 minutes ([219.4, 276.5], *SD* = 158.3) or 4 hours and 7 minutes. Hence, the 118 participants who filled in the EEOLD for at least one day spent around four hours with extramural English per day, which at first appears to be an incredibly large amount of time given that these teenagers attend school, have to do homework, need some sleep and probably also have hobbies that do not involve English.

Considering that 4 hours and 7 minutes are the mean EE time, there must be students in this sample whose exposure far exceeds 4 hours.[166] For this reason, it is interesting to see how the students themselves evaluate their contact time with English: at the end of the EEOLD participants were asked whether the amount of time spent with English on the day they had just recorded was about the same as always, or more or less time than usual. In total, 176 diaries were classified as a 'usual' amount of exposure by the students, 154 as 'less than usual' and only 27 as 'more than usual'. When only looking at the 176 diary entries which participants classified as the 'usual' amount of time, the average time was 269 minutes ([236.7, 301.8], *SD* = 166.0) or 4 hours and 29 minutes, which is more than the overall mean time of 4 hours and 7 minutes. In contrast the mean time for days on which participants spent less time with EE than usual according to their own judgment was 234 minutes ([199.8, 267.5], *SD* = 165.6) or 3 hours and 53 minutes. Both of these findings support the finding that on average a 15- or 16-year-old Viennese teenager spends around 4 hours a day with English during their leisure time, although it needs to be acknowledged that there is great variation as indicated by the large standard deviation.

Time spent with EE was further analysed in relation to weekdays and weekends since it is plausible that teenagers spend more time with English on weekends when they have more free time at their disposal. The 293 diary entries which recorded EE time for weekdays showed a mean of 247 minutes ([216.1, 277.2], *SD* = 165.8) or 4 hours and 6 minutes, which is very close to the overall mean of 4 hours and 7 minutes, whereas the 90 entries concerning Saturdays and Sundays revealed a mean of 335 minutes ([278.3, 392.6], *SD* = 276.7) or 5 hours and 35 minutes. This result shows that at weekends students did indeed engage in EE activities for longer periods of time and adds to the plausibility of the overall results, even if the overall amount of EE exposure is still surprisingly large.

A final analysis with regard to overall EE time was inspired by discussions concerning the role of music among the participants in the focus group interviews (see section 7.3.3). To explore the hypothesis that listening to music, possibly as a background to other tasks and activities, could account for the astonishing amount of time spent with EE, the calculations for EE mean time were carried out a second time taking all EE activities but music into account. The result shows

---

[166] This is the case although 27 diary entries with improbable or outright impossible estimates of time spent with EE between 12 and 28 hours per day were excluded from the dataset as discussed in section 5.3.5.2.

that EE mean time without music drops to 201 minutes ([176, 225.8], *SD* = 138.0) or 3 hours and 21 minutes, which is exactly 47 minutes below the previous estimate including all activities. While this finding indicates that music is certainly an important part of students' extramural activities, it does not support the hypothesis that music accounts for the largest part of time spent with EE.

In addition to the overall mean time, it is interesting to look at how participants allocate their EE time to different activities. Like the CLISS language diary (Sylvén, personal communication 7.10.2014) on which it was modelled, the EEOLD is organized according to language skills (see section 5.3.3.2). Thus, mean EE time can easily be calculated for each of the four language skills and gaming, which includes several language skills. In accordance with the findings on the frequency of EE activities, the results for EE mean time also show that listening activities are most common with an average of 122 minutes ([106.7, 137.2], *SD* = 84.3) or 2 hours and 1 minute per day. As mentioned above, just below 50 minutes are spent on listening to music, 24 minutes on video clips and another 22 minutes on series. Next are the reading activities with 54 minutes ([46.1, 62.0], *SD* = 84.3) a day; these are split among many different activities, but most time was spent on reading in social media with about 9 minutes and reading books with 8 minutes on average. Speaking activities are in third place when it comes to mean time with 30 minutes ([23.8, 26.3], *SD* = 34.7) a day. Again, this can mostly be explained by singing, which is done for about 20 minutes and thus far longer than any other speaking activity. Multi-skill activities are carried out for an average of 24 minutes ([15.2, 32.2], *SD* = 47.1); here 9 minutes are spent with online and 7 minutes with offline gaming. Writing activities come last with a mean time of 19 minutes ([15.4, 22.9], *SD* = 20.7) per day with none of the individual activities being done for more than 3 minutes on average.

The analysis of the EEOLD data reported in this section indicates that 15- to 16-year-old teenagers in Vienna spend an average of 4 hours and 7 minutes per day with extramural English. While this may be considered an astonishingly large amount of time, particularly if compared to the three hours of English exposure per week at school, more detailed analyses indicate that this result is plausible: first, the language diary entries classified by participants as 'average exposure' show an even higher but roughly comparable figure; second, engagement with EE is longer on weekends than on weekdays; and third, this result is supported by the findings of the qualitative strand as shown in section 7.2.2. A further analysis of time spent with EE according to language skills supports the findings on the predominance of receptive skills in the previous section with approximately two hours spent on listening activities followed by almost one hour of reading. In contrast, participants engage in speaking, writing and multi-skill activities for half an hour or less per day on average.

6.3.3 Reasons for engaging with extramural English

Having established that students spend a substantial amount of time with English during their spare time, the question remains why they choose to do so. Table 6.18 presents the responses to six statements that students were asked to rate in the EEQ (item set 2e). The explanations given in these range from availability, international contacts, and the desire to learn to more aesthetic criteria and personal interests. In addition, there was an 'I have another reason' category at the end of the item, which allowed participants to add further personal reasons if they wanted to.

Why students use English in their spare time	strongly disagree (%)	rather disagree (%)	rather agree (%)	strongly agree (%)	*N*
Many things sound better in English.	7.49	7.49	26.2	58.82	187
I am interested what things (e.g. films, books) are like in the original version.	6.88	20.11	30.16	42.86	189
I would like to improve my English skills.	6.91	13.83	39.36	39.89	188
I just enjoy using English.	9.09	18.72	38.5	33.69	187
Many things are only available in English (at least temporarily).	7.41	21.16	44.44	26.98	189
I need English for international contacts and friendships.	23.53	27.27	22.99	26.2	187

Table 6.18: Reasons for using English in spare time activities

Unexpectedly, the most important reason among those listed in Table 6.18, which nearly 60% of participants say fully applies, is that many things sound better in English. This is further corroborated by the fact that of the 19 responses in the 'other category' five underline that English is better, or sounds funnier or more interesting than other languages and in particular German, and one person argued that English media have better content than German ones. 43% fully agree with the next most popular statement which points to a similar concept: interest in the original language versions. Again, three participants also gave original versions as their personal reason for consuming media in English in the open item. Almost 80% of the students concur or fully concur that their engagement with English during their leisure time also stems from a wish to learn and improve their English skills, but at the same time nearly as many participants (73%) also agree that they just enjoy using English, which points to the conclusion that learning is a welcome by-product for students although many probably do not actively pursue it during their spare time activities. Yet, two participants mentioned improving their English or helping a family member to improve theirs as their personal reasons in the open question. The fact that many things are only available in English until dubbed versions are broadcast or published does not appear to be as decisive a factor because only 27% state that this statement fully applies and 45% partially. Lastly, using English for international contacts and friendships appears to be the least influential reason, but interestingly, participants' opinions are spread widely from not applicable at all to fully applicable with about a quarter choosing each of the four response options. Further reasons given by students in the 'other category' relate to their future career plans or plans to move abroad, to the fact that

communication or information searches in English are perceived as quicker and easier, to communication with tourists, and to the specialized niche activity of programming.

On the whole, aesthetic reasons and the desire to read or watch original versions seem to be the strongest explanation for students' use of English in their free time. It is however worth pointing out that more than half of the participants said that all reasons given in item 2e fully or partially apply, except for the last statement on international contacts. Hence, engagement with extramural English seems to result from an interplay of several motives, which may also include factors not taken into account in this study.

In addition to reasons for using EE, participants were also asked about their English-language favourites among books, games, films and series (EEQ item set 2b). Their responses show an impressive range of different genres and titles: overall, participants named 59 different books, 48 games, 97 films and 92 series. Among both books and films, the *Harry Potter* series is the clearest favourite with 12 and 10 mentions respectively, but participants read a variety of genres including fantasy, crime and popular young adult novels such as *The Fault in our Stars*, *Looking for Alaska*, or *Slam*. *League of Legends* was the most popular game with 14 mentions, followed by several shooter games such as *Counterstrike*, *Overwatch* or *Battlefield* and sports games such as *FIFA* or *NBA2k*. The range of films and series mentioned is particularly varied including, action, fantasy, crime, teenage movies and romance. Especially recent hit series such as *Games of Thrones*, *Prison Break*, *The Walking Dead* or *The Vampire Diaries* also seem to be popular among Viennese teenagers because each received more than ten mentions. In terms of other viewing 115 participants mentioned vloggers and YouTube videos in general, but they also watch more specific genres such as let's plays, tutorials, comedy or news. In comparison, online blogs are much less popular with only 19 mentions, but students also like to read news, posts on social media, fan fiction and online articles in English. While these results certainly do not depict the whole EE environment of Viennese teenagers, they highlight again that the out-of-school engagement with English is highly diverse and strongly depends on individual preferences as discussed in section 6.3.1.

6.3.4 Differences in engagement with extramural English according to influencing factors

So far, engagement with extramural English has been described in isolation, but it is essential to consider it in relation to potential influencing factors (see section 5.3.3.1). This section explores differences in EE activities with regard to gender as well as linguistic and socioeconomic background. First, these will be described in relation to the EE median score indicating overall frequency of engagement and mean time spent with EE before turning to individual EE activities.
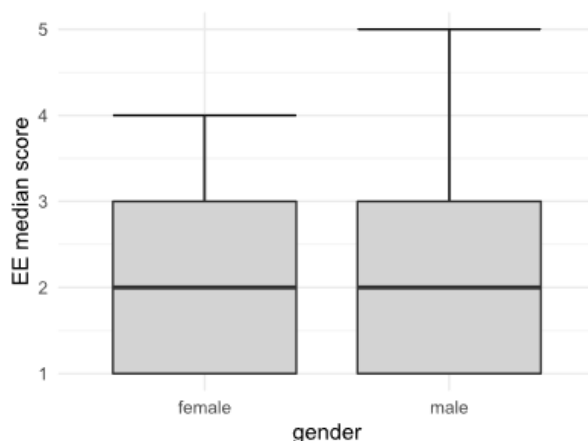
Figure 6.7: Boxplot (left) displaying median and interquartile range for EE median score according to gender (N = 188)

As indicated by Figure 6.7 the median number of EE activities of both girls (Mdn_{girls} = 2 [2,2] n_{girls} = 109) and boys (Mdn_{boys} = 2 [1, 2], n_{boys} = 79) are exactly the same, although the range is higher for boys indicating that at least one of the male participants encounters English more frequently during his leisure time activities. This is mirrored in the number of daily EE activities: in both groups there are participants who report no daily EE activities, but the maximum number of activities for girls is 25, whereas it is 33 activities for boys. [167] The difference between the mean number of daily activities (M_{girls} = 8.94 [7.94, 9.99], SD_{girls} = 5.6, M_{boys} = 10.1 [8.67, 11.73], SD_{boys} = 6.95) is, however, not statistically significant as tested in a Wilcoxon rank-sum test (W = 3967.5, p = .359) and shows a very small effect (r = -.07 [-.21, .08]). [168]
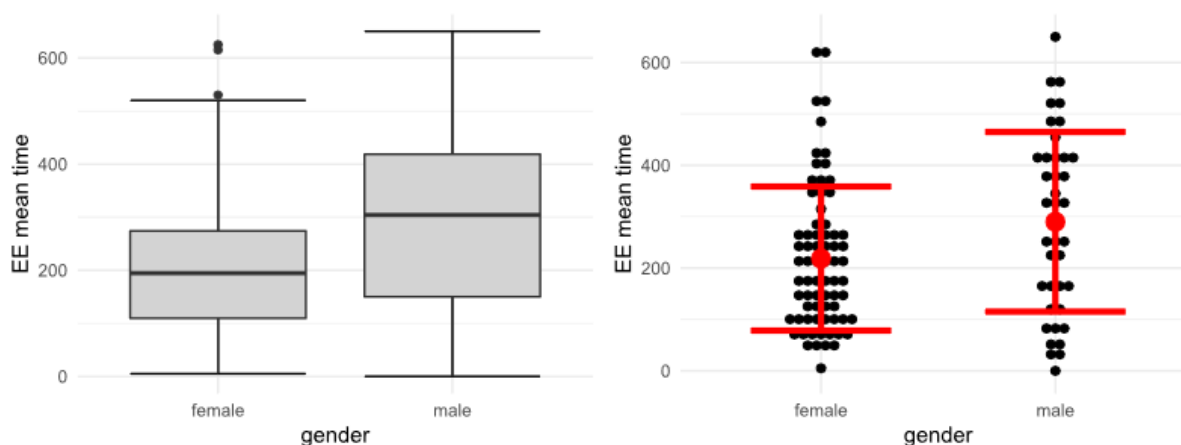
Figure 6.8: Boxplot (left) displaying median and interquartile range and dotplot (right) showing mean and standard deviation (red error bar) of EE mean time according to gender (N = 112)

---

[167] Please note that the number 1 on the vertical axis in Figure 6.7 represents the response option '(almost) never'.
[168] As set out in Chapter 5, all statistical analyses are based on the conventional significance level of α = .05, unless otherwise reported. Correspondingly, 95% confidence intervals (CIs) are reported for all point estimates. In addition, the Pearson correlation coefficient r is used as the common effect size in this thesis to allow for easier comparisons and a simple interpretation of magnitude of effects.

When looking at time spent with EE, the difference between female and male participants is more pronounced as shown by the plots in Figure 6.8. The boxplot on the left indicates that the median is much higher for boys (Mdn_{boys} = 304 [192.5, 386], n_{boys} = 40) than for girls (Mdn_{girls} = 194.5 [145, 242] n_{girls} = 72) and the dotplot on the right also indicates that on average male participants spend more time with EE (M_{boys} = 290.1 [237.3, 344.9], SD_{boys} = 174.9) than their female counterparts (M_{girls} = 218.5 [190.2, 254.6], SD_{girls} = 140.2). In contrast to the number of daily EE activities, the difference in mean time spent with EE according to gender is statistically significant as shown by a Wilcoxon rank-sum test (W = 1101, p = .040). The effect (r = -.19 [-.37, -.01]) remains, however, relatively small.

In addition to gender, several other potential influencing factors are also taken into account; Figure 6.9 displays the relation between these and the frequency of EE exposure summarized by the EE median score. The data are presented both graphically and numerically in a scatterplot matrix; beginning on the left each of these will be discussed in turn. The first scatterplot in the bottom row displays the relation between the EE median score and the SES summary index integrating the highest levels of parental educational attainment and occupational prestige (see section 5.3.5.1). The loess curve indicates that this relationship is positive but not strong, and the corresponding correlation analysis shows a small, but statistically significant association (τ = .18 [.07, .29], p = .003) suggesting that students from more privileged families engage in EE activities more frequently. Related to SES is the number of books available at home and students' access to various media devices, which could also influence their EE behaviour.[169] As can be seen from the respective plots and correlations, the number of books is significantly and relatively strongly related to SES (τ = .39 [.30, .48], p < .001), whereas access to media displays a small, but significant correlation with SES (τ = .22 [.11, .33], p < .001). In relation to the EE median score both the number of books (τ = .14 [.02,.27], p= 0.017) and the availability of different media access points at home (τ = .16 [.00, .27], p = .008) only show small, if significant, effects.

With regard to linguistic background, the relationship between the EE median score and the number of home languages is very small and not statistically significant (τ = .08 [-.06, 20], p = .235), although the graph tentatively suggests that highly multilingual participants speaking four different languages at home engage in more EE activities. It is also interesting to note in this respect that the number of home languages correlates negatively with the SES summary variable (τ = -.30 [-.42, -.18], p < .001) and the two related variables, which means that any conclusions concerning the effect of multilingualism need to be interpreted with caution as they could be mediated by socioeconomic effects.

---

[169] The media access variable refers to access to technical devices available at students' homes (see section 5.3.3.1). It includes owning a smartphone and/or a computer, the availability of internet and access to a television set, a DVD player, a radio, an mp3-player, a tablet, a gaming console and an e-book reader at home.
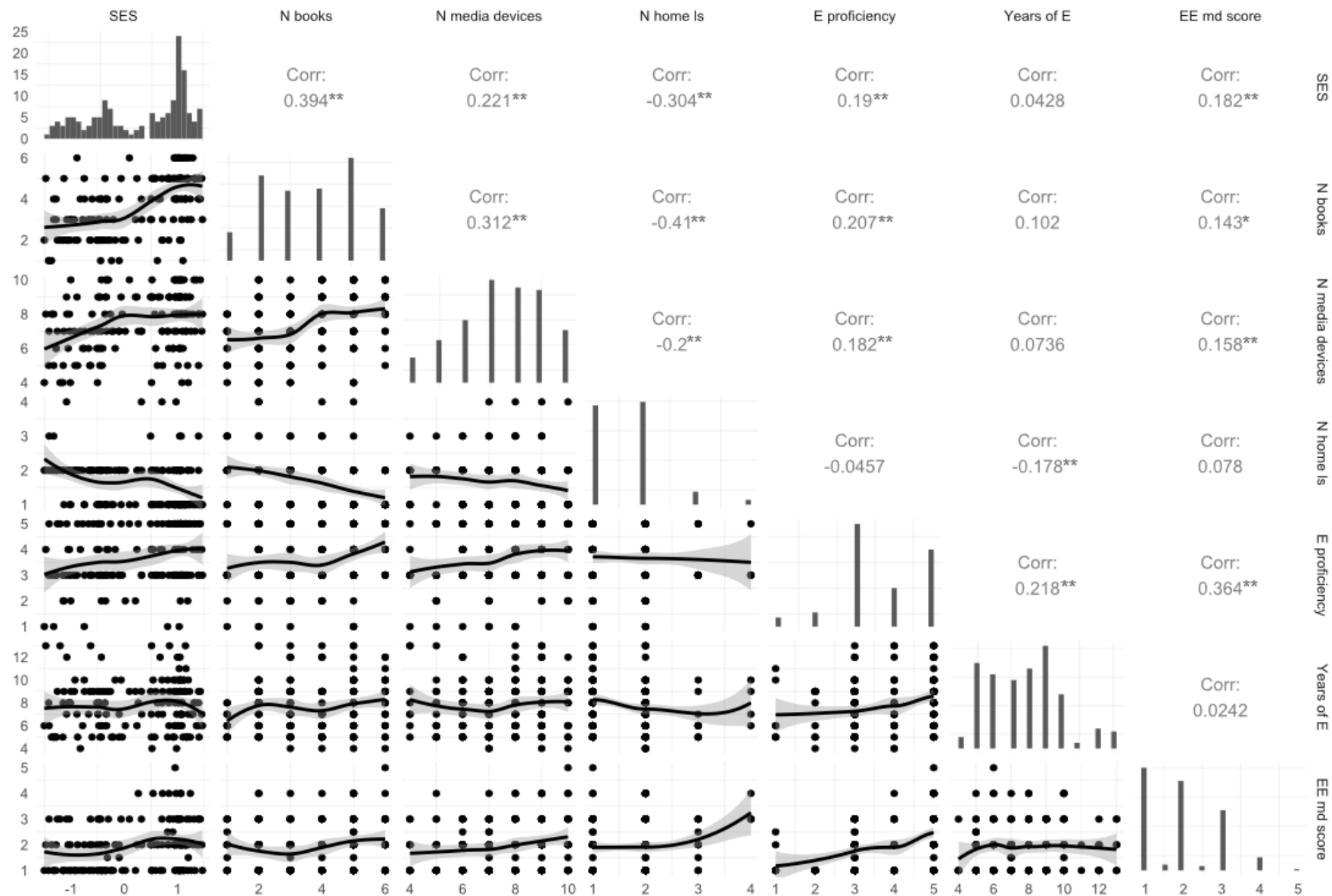
Figure 6.9: Scatterplot matrix displaying relations between influencing factors and median EE score graphically (lower half) and numerically through Kendall's tau (upper half, * p < .05, ** p < .01)

In relation to students' knowledge of English, self-assessed English proficiency correlated significantly and positively with the EE median score (τ = .36 [.25, .47], p < .001) with a medium effect. This finding begs the question of which came first, similar to a 'chicken or egg' dilemma: are more proficient students more likely to engage with extramural English during their leisure time or does increased exposure to English lead to higher proficiency? While the current study is not designed to answer this question, this is a point that certainly warrants further investigation. One aspect that is also of interest in this respect is the relation between the number of years spent learning English and the EE median score. The last scatterplot in the bottom row suggests that this relation is not strong and the correlation analysis confirms that it is not statistically significant (τ = .02 [-.09, .14], p = .686). Hence, a preliminary conclusion could be that for the frequency of engagement with EE it is not important how long students have been studying English, but which level of proficiency they have achieved during this time. This may be an initial indication that there is a two-way relationship between overall language proficiency and extramural English: more proficient students probably engage in EE activities more frequently, and perhaps earlier than their peers, and this in turn may lead to practice and learning effects and thus to an increase in their proficiency.

The influencing variables described above and displayed in Figure 6.9 were also analysed in relation to mean time spent with EE. A corresponding scatterplot matrix can be found in Figure B.1 in Appendix B, while information on confidence intervals and p-values corresponding to the correlations is given in Table B.7. Interestingly, all variables showed smaller correlations with EE mean time than with the EE median score with the exception of length of English instruction: the number of years spent learning English correlates positively with mean time spent with EE (τ = .09 [-.05, .23], p = .175), but the relationship is not statistically significant.

In addition to looking at the relationships between potential influencing factors and the EE median score as a summary variable, it is worth exploring differences for individual EE activities as well. Here, gender, socioeconomic status and self-assessed overall English proficiency were selected as the most interesting factors to investigate. Wilcoxon rank-sum tests were used to analyse gender effects across all individual EE activities; this section describes activities for which the difference between female and male participants is statistically significant at α = 0.05 (see Table B.8 for a complete summary of results). Beginning with those activities that showed the largest effect sizes (see section 5.3.6), it was found that the greatest effect of gender concerned digital games: male participants play significantly more multiplayer online games (W = 941, p < .001, r = -.74 [-.80, -.67]) and computer or console games in English (W = 1191, p < .001, r = -.64 [-.72, -.55]). In combination with gaming they also use VOIP services such as Skype (W = 182, p < .001, r = -.42 [-.59, -.22]) and in-game chats (W = 191.5, p = .001, r = -.39 [-.57, -.18]) more often to communicate with fellow players. Furthermore, boys are more likely to use Skype to speak English (W = 3331, p = .002, r = -.22 [-.36, -.08]), which could potentially also be connected to gaming practices, and they also play English-language games more frequently on portable devices (W = 3002, p < .001, r = -.26 [-.39, -.12]), even if these effects are much smaller.

Concerning activities other than gaming, male participants watch significantly more video clips (W = 3118, p = .001, r = -.24 [-.37, -.01]) and more films on the internet without subtitles (W = 3092, p = .007, r = -.20 [-.34, -.06]) and they are more likely to read comics (W = 3484, p = .009, r = -.19 [-.33, -.05]).

Female participants on the other hand are significantly more likely to engage with music by singing along to English-language songs (W = 5942, p < .001, r = -.37 [.49, -.24]) or singing themselves (W = 5874, p < .001, r = -.33 [-.45, -.19]). They also listen to music more both on CD and on the radio (W = 5153, p = .0089, r = -.20 [-.33, -.05]) or at concerts (W = 5184, p = .004, r = -.21 [-.34, -.07]). In addition, they are more likely to read English-language lyrics (W = 5016, p = .047, r = -.14 [-.28, .00]) or engage with translations of lyrics (W = 5200, p = .009, r = -.19 [-.33, -.05]). Moreover, girls also read more books, both in print (W = 5020, p = .006, r = -.20 [-.34, -.06]) and as e-books (W = 4673, p = .049, r = -.15 [-.28, .00]), more stories (W = 5290, p = .003, r = -.22 [-.35, .07]) and more information print texts, such as recipes or manuals (W = 4913, p = .039, r = -.15 [-.29, -.01]). They are more likely to write stories (W = 4803, p = .020, r = -.17 [-.31, -.03]) or a diary in English (W = 4260, p = .001, r = -.25 [-.38, -.10]), even though not many of them generally do (compare section 6.3.1). Interestingly, girls use social media significantly more than boys in that they read (W = 4921, p = .006, r = -.20 [-.34, -.06]) and write (W = 5090, p = .011, r = -.19 [-.32, -.04]) status updates or comments significantly more often in English, but there is no significant difference with regard to reading and writing messages. Lastly, female participants also watch more films on DVDs, both with (W = 5213, p = .003, r = -.22 [-.36, -.08]) and without subtitles (W = 5155, p = .005, r = -.21 [-.34, -.06]).

While gender differences reach statistical significance at the customary α = 0.05 level for all of these activities, the effects are often very small (see section 5.3.6) and thus their practical significance should not be overestimated. On the whole, the gender differences found for individual EE activities closely resemble those that were found for participants' general leisure time activities described in section 6.1.3: boys play digital games more frequently overall and the same holds true for playing in English, while girls read more often in general and in English and engage more intensively with music.

Relationships between individual EE activities and self-assessed overall English proficiency and SES were investigated using Kendall tau correlations; here only those correlations that show at least small effect according to Plonsky and Oswald's (2014) threshold of r = .25 will be reported, for a complete overview the interested reader is again referred to Appendix B (see Table B.9). The strongest relationship with self-assessed English proficiency was found for thinking or talking to oneself in English (τ = .37 [.26, .48], p < .001) suggesting that participants with a higher level of English engage in 'inner speech' more frequently. Reading English-language books in print form (τ = .33 [.22, .44], p < .001) and reading information texts online (τ = .32 [.20, .43], p < .001) showed small to medium effects in relation to proficiency, as did the use of search engines in English (τ = .32 [.21, .43], p < .001). Online communication in form of speaking English on Skype or similar services (τ = .29 [.17, .41], p < .001), chatting in English (τ = .25 [.12, .38], p = .001) as

well as reading e-mails (τ = .26 [.15, .38], p < .001) and writing them (τ = .29 [.17, .40], p < .001), was also positively related to proficiency, as was writing notes and lists in English (τ = .27 [.14, .38], p < .001). Six of these nine EE activities found to show at least a small correlation with overall proficiency are of an interactive nature and/or entail productive language use. However, as we have seen in section 6.3.1, with the exception of using search engines none of these activities are among the most popular. With respect to the socioeconomic status of participant's families none of the individual EE activities reached an effect equal to or greater than τ = .25, which indicates that SES does not have practical effects on the frequency of individual EE activities.

Summarizing the main findings with regard to the types and amount of contact with extramural English, we have seen that there are nine common activities which more than 50% of the participants engage in at least a few times a week. Listening to music, watching online video clips and reading in social media are the three most popular activities with over 75% reporting to do these at least a few times a week and over 50% almost every day. At the same time, the analysis has shown that there is a wide range of EE activities that Viennese adolescents engage in: each of the 64 activities listed is done at least a few times a month by at least one of the participants and some participants engage in additional 'niche activities'. This finding indicates that in addition to a few highly popular EE activities, teenagers have highly individualized EE environments in line with their general interests and leisure time preferences. The majority of the most common EE activities are carried out online and involve language in a receptive way, listening is the skill most frequently used followed by reading, while the productive skills of speaking and writing are rarely used in informal out-of-school contexts.

Calculation of the mean time spent with EE based on the language diary data suggests that on average Viennese 10th-grade students spend approximately four hours a day with English outside school. This result may be surprising for several reasons, but further analyses in relation to weekdays and weekends, the role of music, and most importantly, participants' evaluation of this outcome in the qualitative strand (see section 7.2.2) indicate that it is plausible. Questionnaire data show that the main reasons for this extensive engagement with EE are the perception that many things sound better in English and an interest in original versions. Interestingly, the majority of students agreed both with the statement that they would like to improve their English, and the statement that they simply enjoy using English in their spare time. Availability and international communication also appear to play a role as motivations for EE activities, but analysis of the qualitative data (see section 7.2.3) sheds further light on this issue.

Finally, the relationships between EE and several influencing factors were also explored: frequency of engagement with EE showed the strongest correlation with overall self-assessed English proficiency followed by SES and the two SES-related variables of the number of books available at students' homes and media access. Significant gender differences were found with regard to mean time spent with EE, but not for overall frequency of engagement. In contrast, length of English instruction and a multilingual language background do not appear to play a

role. Individual EE activities show the same gender difference as general leisure time activities (see section 6.1.3) and for overall English proficiency nine individual activities showed correlations greater than .25, six of which involve productive use of language. Interestingly, none of the individual EE activities correlated significantly with SES at a level of τ = .25.

6.4 Extramural English and vocabulary size

6.4.1 Receptive vocabulary size

This section presents the results with regard to receptive vocabulary size as measured by V_YesNo (Meara 2015a). As discussed in Chapter 5, there are several ways of scoring Yes/No tests; therefore, methodological issues are further explored here. First, the influence of test takers' characteristics on guessing and the number of false alarms is investigated. Second, the results of the scoring formula are compared to students' performance on 20 translation items they filled in immediately after taking the V_YesNo test in order to investigate in how far the checklist format accurately assesses meaning recall (see section 5.3.5.3) Third, the scores produced by these two scoring formulae are also briefly compared to previously suggested formulae.

Basic descriptive statistics for the V_YesNo data were given in section 5.3.5.3; these show that in comparison to other studies the FA rate was relatively low, but 25 participants still had to be excluded because they exceeded a reliability threshold of 15 FAs. It is therefore of interest to analyse whether any specific participant characteristics influence the FA rate because, as mentioned in section 5.3.3.3, the number of FAs may not only reflect the students' confidence in their knowledge, but also character traits or background factors (Eyckmans 2004). For this reason, the FA rate in the V_YesNo data was analysed in comparison to the three participant variables gender, socioeconomic background and overall language proficiency.[170]

First of all, the relationship between the FA rate and the hit rate is of interest: Figure 6.10 suggests that there is a positive correlation since many of the participants who scored a large number of hits, also had a large number of false alarms. This is confirmed by a Spearman rank order correlation (r_s = .45 [.31, .57], p < .001, R^2 = .20), which shows however that the relationship is only of medium strength. A closer look at Figure 6.10 indicates several response behaviours, ranging from students who were very careful not to produce any false alarms in the left bottom corner, over those who produced many hits and false alarms in the right upper corner, to those who produced a large number of hits without hardly any false alarms in the right bottom corner.

---

[170] Please note that for these analyses it was only possible to use those participants who also filled in the questionnaire. Due to absences in one of the two sessions or missing responses, only data from 164 participants could be taken into account.
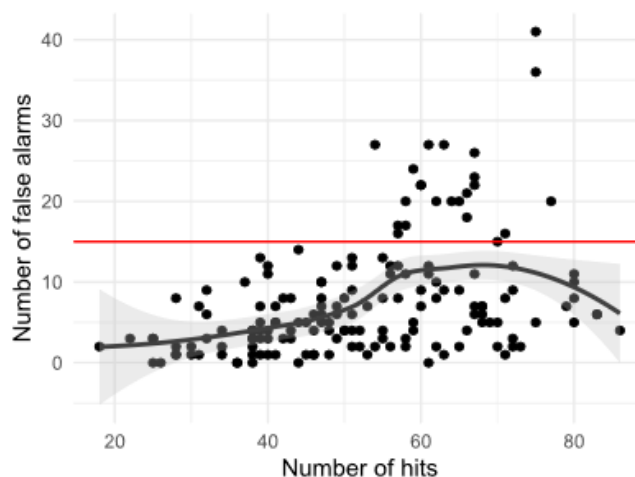
Figure 6.10: Relation between FA rate and hit rate in the V_YesNo data (N = 164). The red line represents the lenient reliability threshold of 15 FAs

Furthermore, the FA rates in the V_YesNo data were also analysed with regard to gender, socioeconomic background and overall language proficiency. Figure 6.11 builds on Figure 6.10 by grouping the data according to gender; it indicates that the lowest numbers of hits were produced by girls, but overall the hit and FA rates of female and male students are very similar. This hypothesis is supported by a Wilcoxon rank-sum test, which showed that any differences were not statistically significant (W = 3244.5, p = .851, r = -.01 [-.17, .14]). Likewise, Figure 6.12, which displays the numbers of hits versus false alarms grouped by socioeconomic status, does not show any conspicuous differences between a high, average and low SES group, which was confirmed by comparison of mean ranks in a Kruskal-Wallis test (H = 0.78, df = 2, p = .678).[171] Finally, Figure 6.13 presents the same plot grouped by self-assessed language proficiency in English. Here, the loess curves suggest that students who rated their overall language proficiency at level B2 produced fewer false alarms in the V_YesNo test than their colleagues who rated themselves at lower levels. However, this relationship between false alarm rate and overall language proficiency did not emerge as statistically significant in a Kruskal-Wallis test either (H = 1.80, df = 2, p = .407). To sum up, the background factors gender, SES and self-assessed English proficiency do not show a statistically significant relation with false alarm rates on the V_YesNo test; hence, test takers' tendency to overestimate their vocabulary knowledge by selecting pseudowords as known words does not follow a pattern related to these factors.

---

[171] Unfortunately, the test statistic H of Kruskal-Wallis-test cannot be readily converted into an effect size r as H has a chi-square distribution with more than one degree of freedom (Field, Miles & Field 2012: 685). For this reason, no separate effect size is given here.
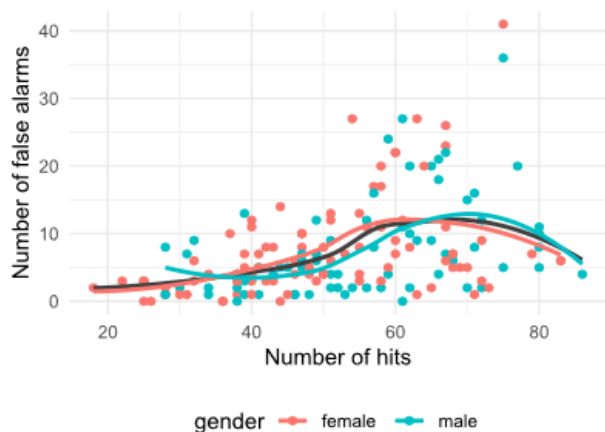
Figure 6.11: Relation between FA rate and hit rate in the V_YesNo data grouped by gender (N = 163)
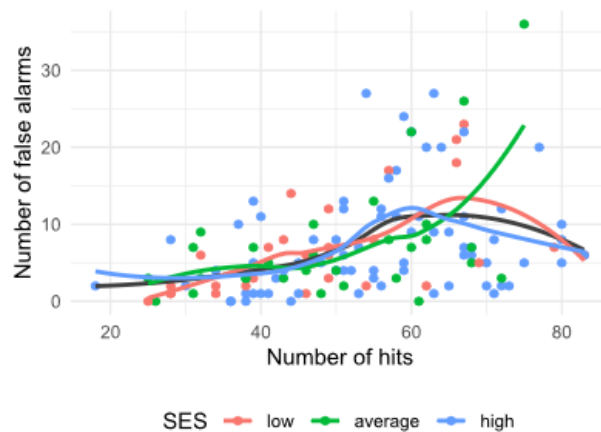
Figure 6.12 Relation between FA rate and hit rate in the V_YesNo data grouped by socioeconomic status (N = 136)
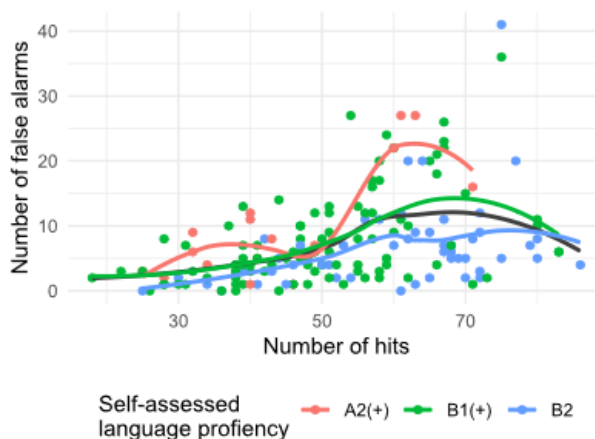
Figure 6.13: Relation between FA rate and hit rate in the V_YesNo data grouped by self-assessed overall English proficiency (N = 164)[172]

As discussed in sections 5.3.3.3 and 5.3.5.3, Yes/No tests can be scored in several ways, but it is not entirely clear in how far the scores of Yes/No tests actually reflect students' knowledge of the form-meaning link of the target items as the test format does not provide any evidence. In the present study the translation task administered immediately after the completion of V_YesNo may shed light on the issue whether participants actually knew the meaning of the words they ticked as known. In addition to scoring V_YesNo with the S-shaped logistic weighting function described in Meara and Miralpeix (2017), the proportion of correct judgements, i.e. the proportion of responses on the V_YesNo test and in the translation task which show the same level of knowledge, can also be used as an alternative to the V_YesNo scoring formula (see section 5.3.5.3). The rationale behind this approach is that if participants overestimated their knowledge on the V_YesNo test in comparison to the translation task for one fifth of the target items, this likely happened for the remaining target words, too and an adjustment of the raw

---

[172] The plus sign following A2(+) and B1(+) indicates that students who rated their overall English proficiency as lying between A2 and B1 or between B1 and B2 were included in the lower category.

hits by proportion of correct judgements can be used to address this issue. Table 6.19 displays summary statistics for this new score (*h×CJ%*) using the proportion of correct judgements in relation to the V_YesNo score based on the logistic weighting function suggested by Meara and Miralpeix (2017).[173] It shows that *h×CJ%* generally leads to a stricter modification and thus lower scores; this is the case when all tests are taken into account and when concentrating on the more reliable sample of tests below the reliability threshold of 15 false alarms. A Wilcoxon signed-rank test shows that the difference between the two sets of scores is statistically significant both for all tests (W = 15014, p < .001, r = -.60 [-.66,-.52]) and for those with fewer than 15 FAs (W = 11133, p < .001, r = -.61 [-.70, -.50]). This suggests that the scores produced by the logistic weighting function significantly overestimate student's knowledge of the form-meaning link in comparison to raw hit rate adjusted by the proportion of correct judgements.

	N	*Min*	*Max*	*Mdn*	95% CI	*M*	95% CI	*SD*
all tests								
h×CJ%	174	1440	7505	3670	[3360, 3896]	3878.05	[3690, 4084]	1303.57
V_YesNo	174	1600	8447	4644	[4454, 4909]	4830.47	[4635, 5048]	1414.78
test <15 FAs								
h×CJ%	149	1440	7505	3450	[3225, 3750]	3846.88	[3635, 4057]	1346.80
V_YesNo	149	1600	8447	4560	[4307, 4971]	4807.26	[4590, 5053]	1498.8

h×CJ% stands for hits (*h*) adjusted by the proportion (%) of correct judgements (*CJ*)
Table 6.19: Summary statistics for h×CJ% scores in comparison to the V_YesNo score

At the same time, the scatterplots presented in Figure 6.14 and Figure 6.15 demonstrate that the two sets of score are closely related. Indeed, the correlation between the V_YesNo score based on the logistic weighting function and the raw hit rate adjusted by the proportion of correct judgements is very strong and positive for all tests (r_s = .89 [.84, .92], p < .001, N = 174) and even higher when only considering those tests with fewer than 15 FAs (r_s = .90 [.85, .94], p < .001, N = 149). Thus, despite the fact that the median difference between the pairs of scores is statistically significant as shown by the Wilcoxon signed-rank test, the Spearman rank order correlation indicates that there is a statistically significant monotone relationship between the two sets of scores in the population, which is of great strength.

---

[173] Please note that the label *V_YesNo score* always indicates the figures calculated with the help of the S-shaped logistic weighting function proposed by Meara and Miralpeix (2017), whereas *h×CJ%* refers to the new scoring method based on the adjustment by correct judgements in relation to the translation items.
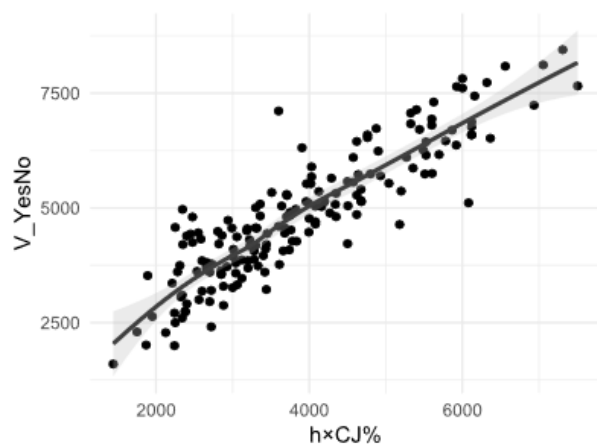
Figure 6.14: Scatterplot of the relationship between V_YesNo scores and h×CJ% scores using all samples (N = 174)
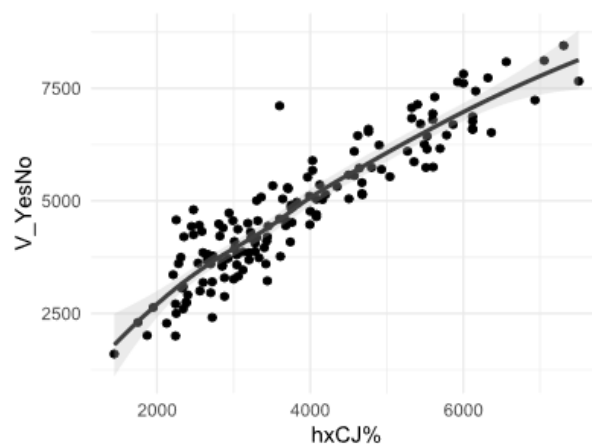
Figure 6.15: Scatterplot of the relationship between V_YesNo scores and h×CJ% scores using reliable samples with FAs <15 (N = 149)

Consequently, it appears that the two sets of scores based on the different methods of adjustment behave similarly for each participant, but that the logistic weighting formula used to calculate the V_YesNo score consistently overestimates participants' knowledge of target items in comparison to the hits adjusted by proportion of correct judgements. This conclusion is also supported by Figure 6.16, which shows similar trends for the two sets of scores overall, but also that the red line representing the V_YesNo scores generally displays higher values.
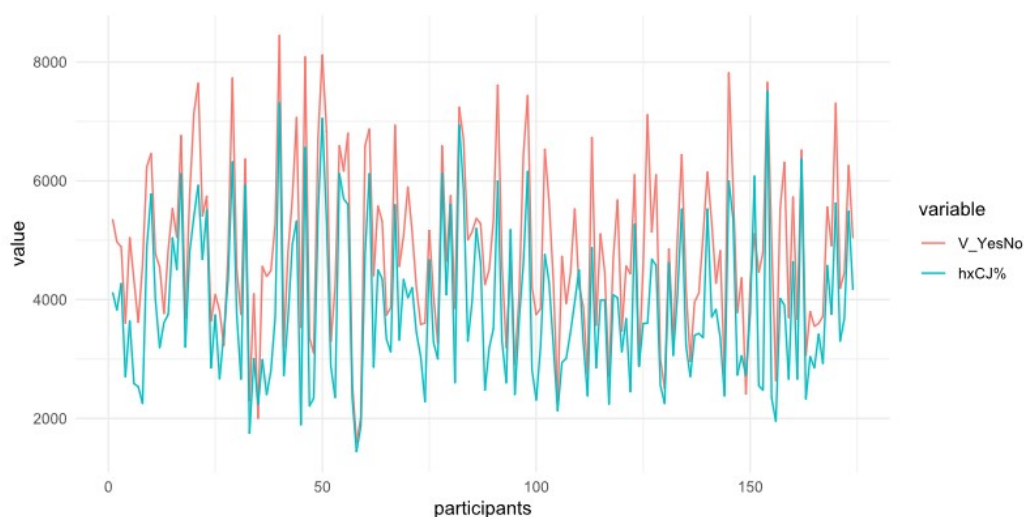
Figure 6.16: Line graph comparing the V_YesNo score and h×CJ% scores for all participants (N = 174)

As discussed in section 5.3.3.3, several methods of adjusting scores of Yes/No tests have been suggested over the years, but since the S-shaped logistic weighting function by Meara and Miralpeix (2017) has only been recently proposed, its relation to the other formulae has not yet been investigated. Following the example of Huibregtse, Admiraal and Meara (2002) and others, a comparison of the V_YesNo scores based on the S-shaped logistic weighting function and the *h×CJ%* score based on the proportion of correct judgements in relation to the translation task is presented in Appendix B (Table B.10 and Figure B.2). This analysis indicates that *h-f* seems to approximate the scores of the logistic weighting function best, even if it is slightly lower.

However, *h×CJ%,* the only adjustment method based on evidence for test takers' knowledge of a sample of target items, consistently produces lower scores than all scoring formulae but Δ*m.* Hence, participants' tendency to overestimate their knowledge of target items does not hold only for the logistic weighting function, but also for *h, h-f, cfg,* and I_{SDT}, while Δ*m* appears to reduce participants' scores more than is justified by a comparison to *h×CJ%.* While certainly worthwhile, a further discussion of the scoring issues related to Yes/No tests is beyond the scope of this thesis and the interested reader is referred to Eyckmans (2004) and other studies discussed in section 5.3.3.3. What can however be concluded at this point is that problem of which scoring formula to use is far from solved and the question of how well scores on Yes/No tests actually reflect knowledge of target items could be an interesting topic for future research in vocabulary testing.

This section described the results of the receptive vocabulary test V_YesNo and discussed methodological aspects related to their calculation. First, the potential influence of gender, socioeconomic background and overall language proficiency on guessing behaviour was analysed, but no statistically significant differences could be found. Second, the results of V_YesNo based on the S-shaped logistic weighting function proposed by Meara and Miralpeix (2017) were compared to several other scoring formulae and in particular to another method of adjusting the V_YesNo data using the number of correct judgements based on a translation task (*h×CJ%*). These comparisons showed that the results of the logistic weighting function are comparable to those of the simpler *h-f* and correction for guessing formulae. However, all of these formulae overestimate the vocabulary knowledge of 15/16-year-old Viennese learners of English considerably in comparison to the number of hits adjusted by the proportion of correct judgements: based on the results of the logistic weighting function participants know approximately 4,800 words receptively on average, while according to *h×CJ%* the mean number of words for which participants could recall the meaning was just below 3,900 words, which is considerably less. This difference is not only interesting from a methodological and interpretative point of view but could potentially result in varying outcomes in statistical analyses. Consequently, all analyses described in the subsequent sections of this chapter were run on both sets of scores, but for reasons of space the analyses based on the first set, the V_YesNo scores using the S-shaped logistic weighting function, will primarily be reported. To allow for comparisons, visualizations and results of all analyses based on *h×CJ%* can be found in Appendix B (Figure B.2 and Table B.10). In general, the findings do however show very similar trends owing to the strong correlation between the two sets of scores.

6.4.2 Differences in receptive vocabulary size in relation to extramural English and other influencing factors

While the previous section was concerned with the overall vocabulary size of Viennese students attending 10th grade of an academically oriented secondary school (AHS), this section goes on to explore factors that potentially have an impact on participants' receptive vocabulary size before giving a detailed description of the connection between EE and V_YesNo scores.
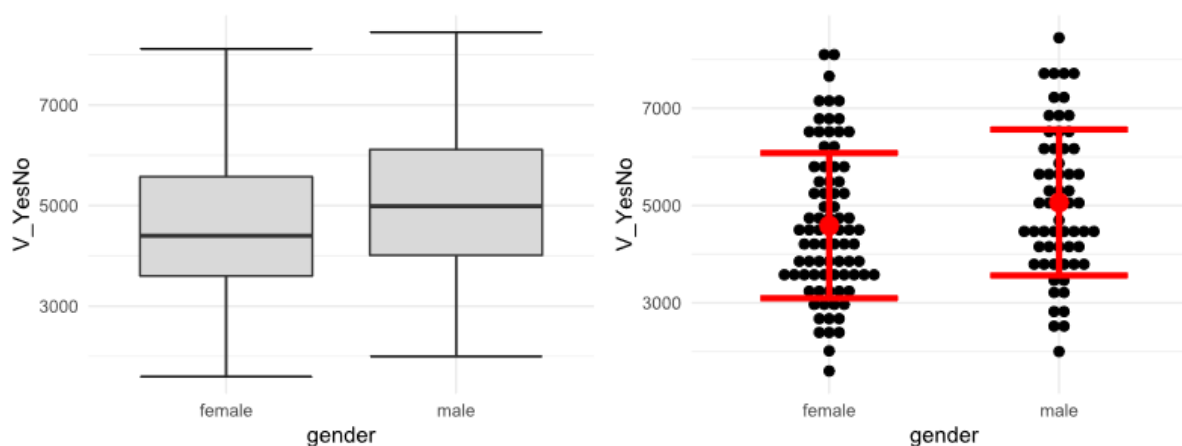
Figure 6.17: Boxplot (left) displaying median and interquartile range and dotplot (right) showing mean and standard deviation (red error bar) of V_YesNo scores according to gender (N = 141)

The first factor taken into account is gender: Figure 6.17 presents a boxplot and a dotplot of the V_YesNo scores for female and male students.[174] As can be seen, the median score for boys (Mdn_{boys} = 4988 [4499, 5537], n_{boys} = 60) is higher than for girls (Mdn_{girls} = 4400 [3879, 4730], n_{girls} = 81) and the same is true for mean scores (M_{boys} = 5063 [4700, 5451], SD_{boys} = 1500.46, M_{girls} = 4588 [4279, 4936], SD_{girls} = 1492.46). Since the test scores of both boys and girls are approximately normally distributed, a t-test was conducted: it shows that the difference in vocabulary size of female and male students (t = -1.86, p = .065) was not statistically significant at α = 0.05 and the effect was small (r = 0.16 [-.01, .31]).

The relationships between the V_YesNo scores and six potential influencing factors are displayed in Figure 6.18.[175] Since the relationships between the different background variables and the EE median score as well as those among the background variables have already been described in section 6.3.4, the focus here is exclusively on the relation of these factors with receptive vocabulary size, and thus on the bottom row and the rightmost column of Figure 6.18.[176] From the first scatterplot showing the relation between the SES index and the V_YesNo score, we can see that there is a positive relationship and a correlation analysis shows that it is statistically significant with a small effect (τ = .20 [.09, .32], p = .001). The two other variables relating to students' socioeconomic background, the number of books (τ = .22 [.11, .33], p < .001) and access to different media (τ = .21 [.09, .32], p < .001), also show small positive correlations with receptive vocabulary size.

---

[174] For results using $h \times CJ\%$ please see Figure B.3 and Table B.11 in Appendix B.

[175] For results using $h \times CJ\%$ please see Table B.12 and Figure B.4 in Appendix B.

[176] The correlations reported in Figure 6.18 correspond to those in Figure 6.9 in section 6.3.4. In Figure 6.18, the V_YesNo score was added as a new variable, whereas overall self-assessed English proficiency is only included in the analysis for EE but not for vocabulary size since vocabulary size can be regarded as an approximation of language proficiency on its own.
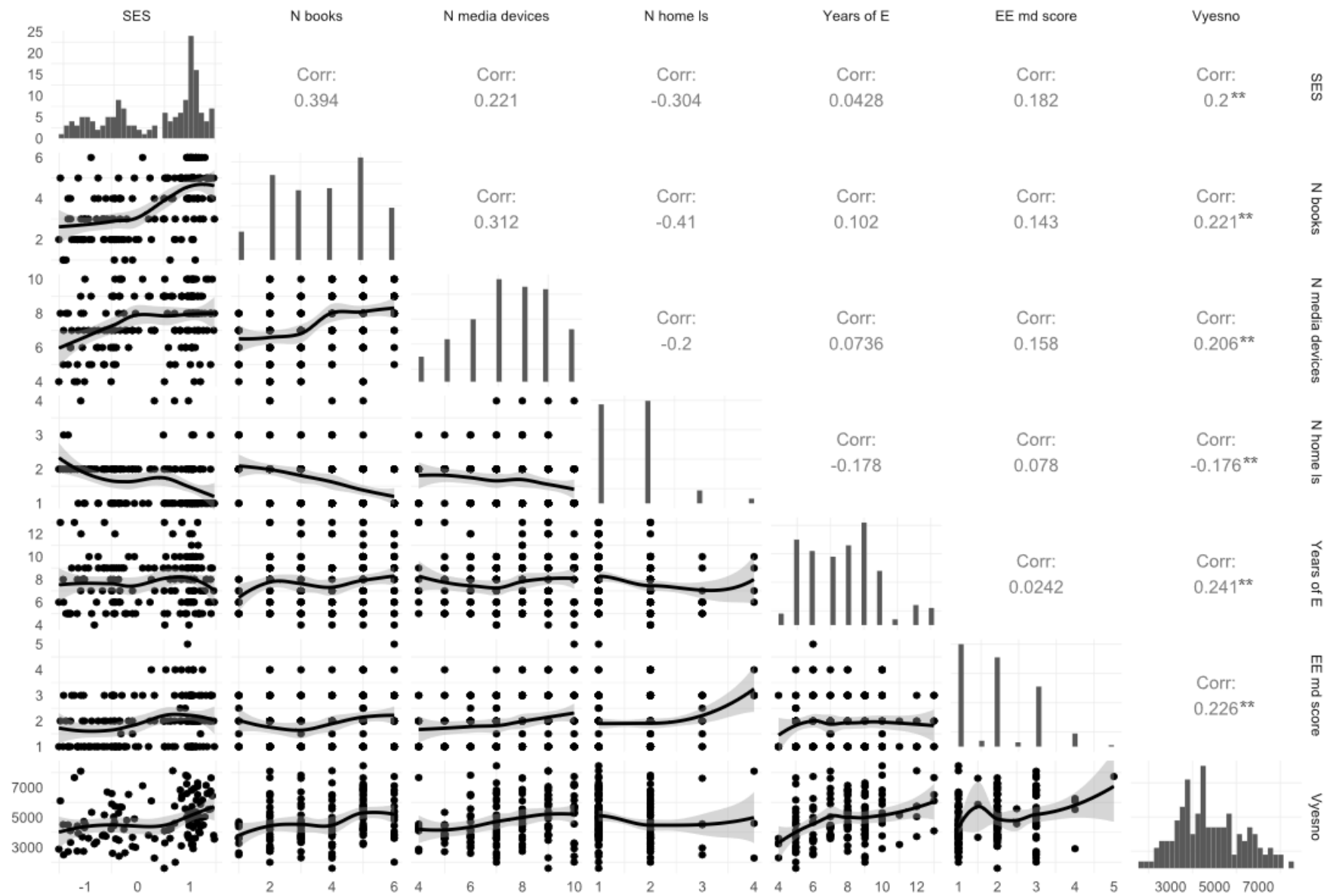
Figure 6.18: Scatterplot matrix showing relations between influencing factors and V_YesNo score graphically (lower half) and numerically through Kendall's tau (upper half, * p < .05, ** p < .01)

Moving on to variables related to the participants' language background, the scatterplot displays a slight decrease in test scores as the number of home languages increases. According to the correlational analysis this slightly negative association between receptive vocabulary size and the number of home languages is indeed significant, albeit with a small effect (τ = -.18 [-.32, -.04], p = .009). This outcome is slightly surprising, but the fact that that the number of home languages and SES show inverse relationships with the V_YesNo score could point to an explanation: as described in section 6.3.4, there is a significant negative correlation between the number of home languages and the SES summary variable (τ = -.30 [-.42, -.18], p < .001). Hence, a possible interpretation could be that the observed negative relationship between the number of home languages and receptive vocabulary size is actually due to a mediating effect of a lower socioeconomic status among multilingual families. To investigate this issue, a partial correlation of the number of home languages and the V_YesNo score with the SES summary variable held constant was calculated (see section 5.3.6). It shows that the negative relationship is weaker and not statistically significant (τ = -.12, p = .058), which supports the interpretation that knowledge of several languages per se is not negatively related to receptive vocabulary size.[177]

Next, length of instruction operationalized as the number of years spent learning English also correlates significantly with receptive vocabulary size (τ = .24 [.13, .36], p < .001). This result is not unexpected, as it is to be anticipated that the length of time spent learning a language has a positive impact on vocabulary knowledge; however, the effect is small. Finally, one of the main questions of this study is whether a positive relationship similar to the one found for length of instruction also exists for additional language input through extramural English. The data show that there is indeed a significant positive correlation between the EE median score and the V_YesNo test score (τ = .23 [.10, .35], p < .001) with an effect size similar to length of instruction. In order to zoom in more closely on the relation between receptive vocabulary size and extramural English, it is informative to investigate whether the scores on the V_YesNo test differ for groups exhibiting different EE behaviours. Since the inclusion of a control group in the research design was not practically possible (see also section 8.1), a comparison of extreme groups can be used to further explore the relationship between out-of-school engagement with English and receptive vocabulary knowledge.

Extreme EE groups were created based on the EE median score: since the average median score was 1.96, all participants with an EE median score of 1 were put in a low EE group (n = 70) and all those with a score equal to or greater than 3 were categorized as the high EE group (n = 51), while the remaining participants (n = 68) were classified as average. Figure 6.19 indicates that both the median and mean scores of the V_Yesno test are indeed higher in the high EE group (Mdn_{high} = 5422 [4600, 6260], M_{high} = 5311.33 [4757, 5856], SD_{high} = 1707.96, n_{high} = 36) than in the average EE group ($Mdn_{av.}$ = 4648 [4314, 5152], $M_{av.}$ = 4951.10 [4597, 5335], $SD_{av.}$ = 1353.93,

---

[177] The computation of the partial correlation was carried out with R package *ppcor* (Kim 2015) and also used Kendall's τ as a measure of correlation. Due to computational issues 95% confidence intervals for partial correlations using Kendall's τ cannot readily be supplied.

$n_{av.}$ = 51) and the low EE group (Mdn_{low} = 4098 [3744, 4460], M_{low} = 4305.02 [3996, 4701], SD_{low} = 1368.26, n_{low} = 55).[178] A Kruskal-Wallis test (H = 11.93, df = 2, p = .003) shows that the differences in receptive vocabulary size between the three EE groups are statistically significant. Post-hoc comparisons with one-tailed pairwise Wilcoxon rank-sum tests using the Bonferroni correction indicate that both the differences between the low EE group and the high EE group (p = .005) and the low EE group and the average EE group (p = .011) are statistically significant. The difference between the average and the high EE group is not statistically significant (p = .294).[179]
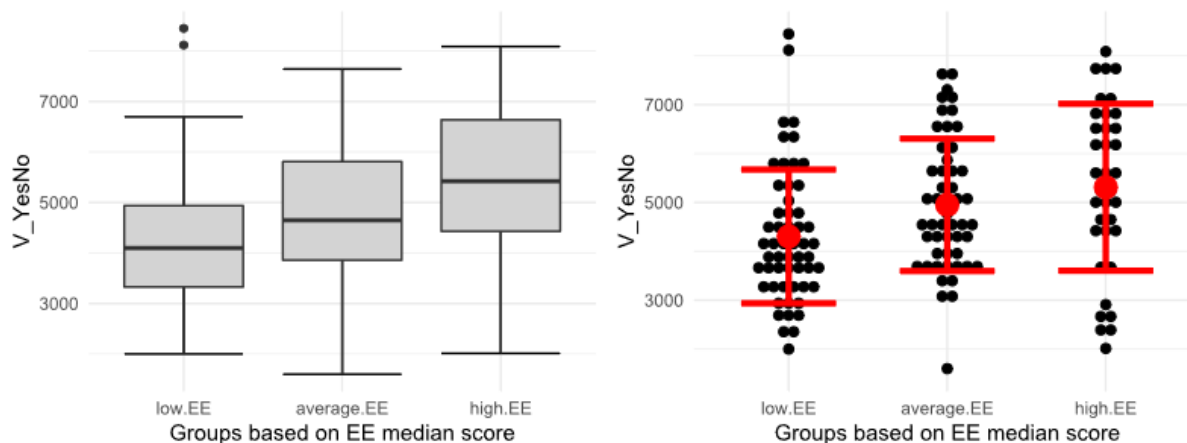
Figure 6.19: Boxplot (left) displaying median and interquartile range and dotplot (right) showing mean and standard deviation (red error bar) of V_YesNo scores according to EE extreme groups based on EE median score (N = 142)

To explore whether length of exposure to extramural English had a comparable effect to frequency of engagement, extreme groups were also created based on the estimates of EE time in the EEOLD, although fewer participants had responded to the online language diary (see section 5.3.4.2). Bearing in mind that the mean time spent with EE per day was 248 minutes (see section 6.3.2), all students who reported spending less than 120 minutes or two hours with EE on average were put in a low EE group (n = 31) and those whose mean exposure time exceeded 360 minutes or 6 hours were regarded as the high EE group (n = 28), leaving 54 students in the average group in between. As can be seen from Figure 6.20, the V_YesNo score rises with the amount of EE exposure similarly to before (Mdn_{low} = 4210 [3672, 4458], M_{low} = 4068.85 [3727, 4441], SD_{low} = 948.90, n_{low} = 26; $Mdn_{av.}$ = 4786.5 [4187, 5327], $M_{av.}$ = 4770.11 [4394, 5156], $SD_{av.}$ = 1322.198, $n_{av.}$ = 44; Mdn_{high} = 5537 [4370, 6100], M_{high} = 5252.14 [4424, 6042], SD_{high} = 1862.98, n_{high} = 21). Again, a Kruskal-Wallis test (H = 8.03, df = 2, p = .018) indicates that there are statistically significant differences in the receptive vocabulary scores of the three groups based on mean time spent with EE. Post-hoc comparisons with one-tailed pairwise Wilcoxon rank-sum tests using the Bonferroni correction show that the difference between the low EE group and the high EE group (p = .010) is statistically significant, whereas the differences between the low EE

---

[178] Data are missing due to participants being absent in the data collection session and the threshold of 15 FAs for V_YesNo tests in this and further analyses.

[179] For results using $h \times CJ\%$ please see Figures B.5 and B.6 as well as Table B.13 in Appendix B.

group and the average EE group (p = .056) and the difference between the average EE group and the high EE group are not statistically significant (p = .383).

Figure 6.20: Boxplot (left) displaying median and interquartile range and dotplot (right) showing mean and standard deviation (red error bar) of V_YesNo scores according to EE extreme groups based on EE mean time (N = 91)

One final aspect to be investigated with regard to receptive vocabulary and EE is the role of niche activities which often require more intense engagement with English and more language production than the most popular activities (see section 6.3.1). For this reason, it is interesting to explore whether the participants engaging in niche activities have a larger vocabulary due to their more intensive contact with English.[180]

A total of 43 participants engage in niche activities at least a few times a month or more often. When comparing the receptive vocabulary size of this sub-sample to the remaining participants, we find that the mean vocabulary size of participants reporting regular niche activities is indeed higher than that of the remaining participants: mean vocabulary size in the sub-sample is 5219 words ([4624, 5780], *SD* = 1559, *n* = 29) based on the V_YesNo score compared to a mean vocabulary size of 4683 words ([4424, 4969], *SD* = 14767, *n* = 113) among those not engaging in niche activities. The dotplots in Figure 6.21 graphically display these data and a comparison of the median score in the sub-sample (*Mdn* = 4971, [4385, 5678]) and the remaining participants (*Mdn* = 4683, [4108, 4735]) shows the same picture.[181] A one-tailed Wilcoxon rank-sum test (*W* = 2008.5, *p* = .031) indicates that the difference in mean receptive vocabulary size as measured by V_YesNo is statistically significant with a small effect (*r* = -.18 [-.35, -.02]).

---

[180] The niche activities used for this exploration are the eight least frequent EE activities in the present sample as well as other EE activities that participants mentioned in the open question at the end of EEQ item set 2a and thus the same activities that were described in more detail at the end of section 6.3.1.

[181] The same is true for an analysis based on the *h×CJ%* score, please see Table B.14 and Figure B.7 in Appendix B for details.
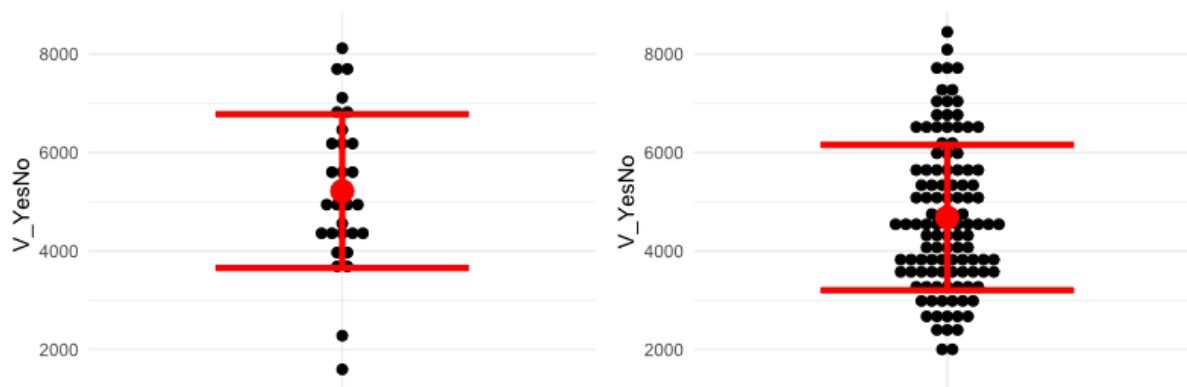
Figure 6.21: Dotplots comparing the V_YesNo scores in a sub-sample of participants engaging in niche EE activities (N = 29, left) and the remaining participants (N = 113, right) showing the mean and standard deviation (red error bar)

The bivariate analysis thus shows that there is a small positive correlation between the EE median score and the V_YesNo score and that the differences between high and low extreme groups are statistically significant based on frequency of engagement and time spent with EE. In addition, a comparison of participants engaging in niche activities and those who do not also shows a statistically significant different in receptive vocabulary size. Parallel analyses using the *h×CJ%* score based on the stricter scoring method mirror these results except for the last comparison, for which the difference was not statistically significant (see Appendix B). Taken together, these findings indicate that there is a positive relationship between EE and receptive vocabulary size, which is further explored using multivariate analysis in the next section.

6.4.3 Modelling receptive vocabulary size

In addition to the bivariate analyses presented in the previous sections, it is desirable to assess the effect of multiple predictors on receptive vocabulary size. Hence, a multivariate model was constructed including several background variables as predictors and the V_YesNo score as the outcome variable. As mentioned in section 5.3.6, I planned to use a linear mixed effects model (e.g. Gałecki & Burzykowski 2013) to take the hierarchical structure of the dataset, which contains individual participants nested in classes and classes nested in schools, into account. However, due to a relatively large amount of missing data the number of cases per group that could be included in the model for receptive vocabulary size was too low for the valid estimation of a mixed effects model. Computational issues occurred in the estimation of random effects because, as shown in Figure 6.22, sample sizes differ vastly between the participating schools and groups. In one school as few as 4 participants remained for inclusion after the application of exclusion criteria and the threshold of 15 FAs in the scoring procedure for V_YesNo.
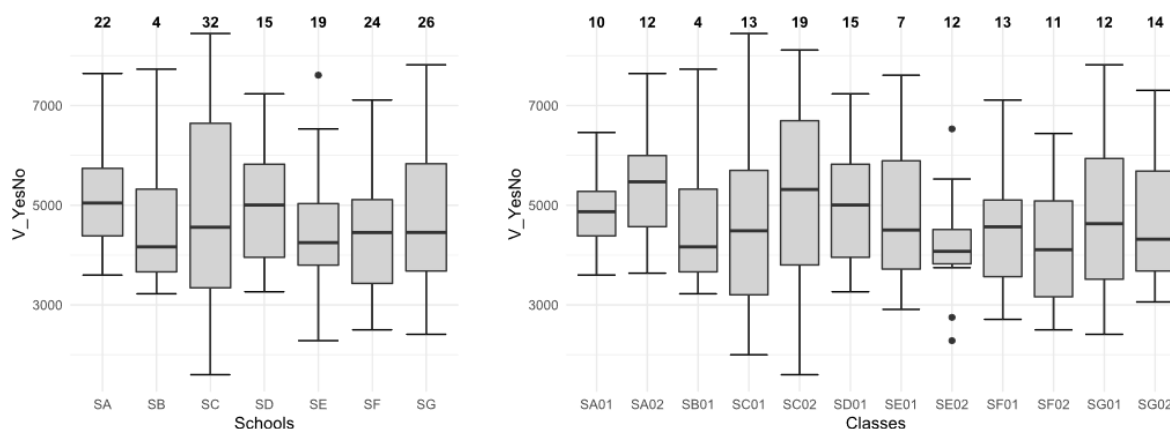
Figure 6.22: Boxplots displaying the V_YesNo scores split by the seven participating schools and twelve participating classes (N=142)

As a consequence, a standard multiple regression model using ordinary least squares estimation was used for multivariate data analysis. Table 6.20 presents the six predictor variables that were taken into account in building the regression model together with their operationalizations. Originally, the median score for the vocabulary learning strategies was also meant to be included as a predictor, but due to the low reliability of the scale (see section 5.3.5.1) it was omitted from the regression analysis. Furthermore, mean time spent with EE could not be included in the model due to the low response rate (see section 5.3.5.2).

Predictor variable	Operationalization
Extramural English (EE)	Median score for extramural English activities
Length of English instruction	Years spent learning English
Socioeconomic status	Summary variable combining highest level of parental education and occupational prestige
Media access	Number of different media devices available at participants' homes
Gender	Gender
Number of home languages	Number of languages spoken at home

Table 6.20: Predictor variables considered for multiple regression model

In terms of sample size needed for a multiple regression model, the present study (*N* = 189) can be deemed sufficiently large (see section 5.3.6), but what complicates the matter is the issue of missing values. As discussed in section 5.3.5.3, the V_YesNo tests of 49 participants had to be excluded from analysis and many of the remaining participants failed to answer one of the EEQ items necessary for the calculation of the summary variables used as predictors. The percentage of missing values across the seven variables (the outcome variable and six predictors) varied between 0 and 25%; in total, 72 out of 189 cases (38%) were incomplete. Since these cases cannot be used in the multiple regression analysis, the sample size was reduced to 117 participants, which is not ideal but not uncommon in empirical research. Multiple imputation by chained equations (Van Buuren 2012; Van Buuren & Groothuis-Oudshoorn 2011) was considered as an option to deal with the issue of missing data, but since multiple imputation for multivariate models is a complex procedure which has a number of assumptions and should ideally include additional variables to predict missing data (Van Buuren 2012), a decision to use listwise deletion and present a complete case analysis was finally reached.

The multiple regression analysis for receptive vocabulary size fit on 117 complete cases resulted in the outcome shown in Table 6.21.[182] The total R^2 value for the model containing the six predictors was 21.3%, meaning that in combination the predictor variables explained 21% of the variance in V_YesNo scores. In addition, the table also presents the squared semi-partial correlation sr^2, a measurement of the relative importance of each variable which "expresses the unique contribution of the IV [independent variable] to the total variance of the DV [dependent variable]" (Tabachnick & Fidell 2013: 144).

	B	95% CI	SE	p	β	sr^2
Intercept	1864.41	[163.33, 3565.49]	858.36	.032*	-0.08	
EE	461.81	[169.66, 753.96]	142.42	.002**	0.28	0.008
Length of English instruction	172.34	[54.17, 290.50]	59.63	.005**	0.25	0.005
SES	163.68	[-161.39, 488.74]	164.03	.321	0.11	0.001
Media access	116.53	[-40.98, 274.05]	79.48	.145	0.13	0.002
Gender (male)	165.68	[-338.83, 670.18]	254.57	.517	0.11	0.000
Number of home languages	-181.43	[614.64, 251.79]	218.60	.408	-0.08	0.000

N = 117, adjusted R^2 = .21 [.12, .41]

Table 6.21: Coefficients of standard multiple regression model for V_YesNo scores

The regression model shows that receptive vocabulary size as measured by V_YesNo is predicted by frequency of engagement with extramural English and the number of years students spent learning English. The effects of the other predictors included in the model are not statistically significant. The standardized $β$ coefficients as well as the values for sr^2 indicate that frequency of engagement with extramural English explains slightly more variance in receptive vocabulary size than the number of years students spent learning English. In general, the very low values for sr^2, and thus for the unique amount of variance explained by the individual predictor variables (1.6% in total), suggest that the largest part of the 21% of variance explained by the model is shared among the predictors.

In the following, diagnostic plots and tests to investigate the assumptions for the regression analysis and the fit of the model are described, which generally suggest that there are no concerns about unmet assumptions and influential data points. The plots in Figure 6.23 show that the relationships between the outcome and predictors variables are reasonably linear so that a multiple linear regression model can be computed. The relation between the V_YesNo score and the number of home languages shows slight curvature but has only very few data points at the right end of the x-axis; therefore, no transformations were applied.

---

[182] Unstandardized coefficients B, p values and 95% confidence intervals are taken from the standard multiple regression analysis, the standardized $β$ coefficients were obtained from a regression analysis with scaled and centred, thus z-transformed, predictors.

Figure 6.23: Scatterplots of the relationship between the continuous predictors included in the regression model and the outcome variable V_YesNo

The assumption of independence is met for both outcome and predictor variables as each value was produced by a different participant. Independence of errors or a lack of autocorrelation was tested using the Durbin-Watson test: the results show a value of 0.05 for autocorrelation and a test statistic of 1.87 with p = .536. According to Field, Miles and Field (2012: 292) the result of the Durbin-Watson test should not be statistically significant and the test statistic should be close to 2; both of which apply to this model. Multicollinearity was assessed using the variance inflation factor (VIF), which ranged between 1.04 and 1.48 for the six predictor variables and was thus well below the critical value of 5 (Larson-Hall 2016: 261). As for assumptions about residuals, Figure 6.24 shows four diagnostic plots: the plot of residuals against predicted values indicates that the homogeneity of variance assumption has not been violated. Next, the normal Q-Q plot shows that the residuals are reasonably close to a normal distribution and the residuals vs. leverage plot makes clear that there are no outliers that exert an undue amount of influence because the dotted line of Cook's distance is not even displayed in the plot. The maximum value of Cook's distance, a measure for the overall influence of a case on the regression model, is 0.24 as shown in the fourth plot, which is well below a suggested critical value of 1 (Field, Miles & Field 2012: 269).

Figure 6.24: Diagnostic plots for the linear regression model: residuals vs fitted values (upper left), normal Q-Q plot of residuals (upper right), residuals vs leverage plot (lower left) and Cook's distance (lower left)

As the model presents one of the core findings of the quantitative strand of this study, the regression analysis was also conducted using the *h×CJ%* score (see sections 5.3.5.3 and 6.4.1) as the dependent variable. Like the model using the V_YesNo score based on the logistic weighting function, this model was based on a complete case analysis with only those V_YesNo tests that had fewer than 15 false alarms (N = 117) and used exactly the same predictors. Table 6.22 presents the parameters of the model using *h×CJ%* as the outcome variable.

	B	95% CI	*SE*	*p*	*β*	*sr²*
Intercept	935.78	[-589.22, 2460.74]	769.51	.227	-0.06	
EE	406.13	[146.44, 665.81]	131.04	.002**	0.28	0.009
Length of English instruction	160.83	[55.99, 265.67]	52.90	.003**	0.26	0.005
SES	202.30	[-87.29, 491.89]	146.13	.169	0.15	0.002
Media access	110.33	[-30.14, 250.80]	70.88	.122	0.14	0.002
Gender (male)	100.45	[-347.22, 548.12]	225.89	.657	0.07	0
Number of home languages	-25.87	[-411.19, 359.45]	194.43	.894	-0.01	0

N = 117, adjusted R^2 = .22 [.13, .42]

Table 6.22: Coefficients of standard multiple regression model for hits adjusted by proportion of correct judgements (h×CJ%) as dependent variable

The model accounts for 22% of variance in receptive vocabulary size, which is slightly higher than the model based on the V_YesNo score. The raw estimates of the coefficients are slightly lower for all predictors except the number of home languages, but again, frequency of engagement with EE and length of English instruction emerge as the only statistically significant predictors. A comparison of the values for sr^2 indicates that the EE predictor and the non-significant SES predictor explain marginally more unique variance, but overall the results of the multiple regression analysis using the stricter scoring method are very similar to the model using the scores based on the logistic weighting function.

In parallel to the model using the V_YesNo scores based on the logistic weighting function, diagnostic analyses for the model using $h \times CJ\%$ as its outcome variable do not show cause for concern. Diagnostic plots, which can be found in Figures B.8 and B.9 in Appendix B, show that the relationships between the $h \times CJ\%$ and the five continuous predictor variables are reasonably linear and that assumptions about residuals have not been violated either. There are no influential data points according to the measure of Cook's distance. The Durbin-Watson test shows a test statistic of 1.85 with p = .384; hence, there is no problem with autocorrelation, and multicollinearity as assessed by the VIF is low for all predictors ranging from 1.03 to 1.37.

6.4.4 Productive vocabulary size

The following sections present the results for productive vocabulary size as measured by Lex30. Basic descriptive information has been included in section 5.3.5.4 together with the details of the scoring procedure. This section briefly analyses the relationship of productive and receptive vocabulary size scores in the present study before exploring the relationships with possible influencing factors and in particular EE in the following sections.

As mentioned in section 5.3.5.4, the mean Lex30 score of the 172 samples that could be used in the analysis is 38.23 ([36.51, 39.92], SD = 11.41), which is about two thirds of a "good native speaker score [... of] about 60 points" (Meara 2009: 136). The median score is 37 [34.5, 40] with a minimum of 7 and a maximum of 69 points. Since there is not enough Lex30 data to allow any further proficiency-related benchmarking, a more meaningful interpretation of these scores is not currently possible, but see section 8.2.2 for a comparison to results of other studies with comparable samples. Nonetheless, the Lex30 scores can be used to rank participants in terms of their knowledge (see sections 3.2.3 and 5.3.3.3).

Since this study includes measures of both receptive and productive vocabulary size, it is of interest to analyse connections between these. Figure 6.25 shows that participating students' Lex30 and V_YesNo scores are positively related and a Spearman rank order correlation confirms that there is a rather strong and statistically significant positive correlation (r_s = .53 [.38, .64], p < .001, R^2 = .28). The relationship in the present study is, however, considerably weaker than the correlation between Lex30 and a Yes/No test (r = .84) found by Meara and Fitzpatrick (2000, see section 5.3.3.3).
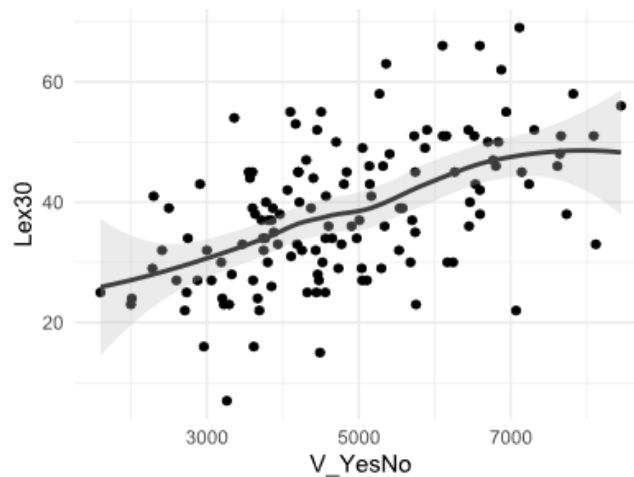
Figure 6.25: Scatterplot of the relationship between productive vocabulary size as measured by Lex30 and receptive vocabulary size as measured by V_YesNo (N = 137)

6.4.5 Differences in productive vocabulary size in relation to extramural English and other influencing factors

This section presents data on the relation between productive vocabulary size as measured by Lex30 and several participant variables analogous to section 6.4.2. Beginning with gender, Figure 6.26 visualizes the Lex30 score for male and female participants and suggests that, similar to the V_YesNo data, boys outperformed girls. The boxplot on the left of Figure 6.26 shows a higher median score for boys (Mdn_{boys} = 38.5 [35.5, 42], n_{boys} = 68) than for girls (Mdn_{girls} = 35.5 [32, 39], n_{girls} = 92) and the dotplot on the right reveals the same for the mean scores (M_{boys} = 40.06, [37.55, 42.62], SD_{boys} = 10.46; M_{girls} = 36.49, [34.11, 39.05], SD_{girls} = 11.81). A Wilcoxon rank-sum test (W = 2552, p = .047, r = -.16 [-.31, .00]) indicates that the difference in productive vocabulary size between male and female participants is statistically significant with boys having higher Lex30 scores; however, the size of the effect is very small.
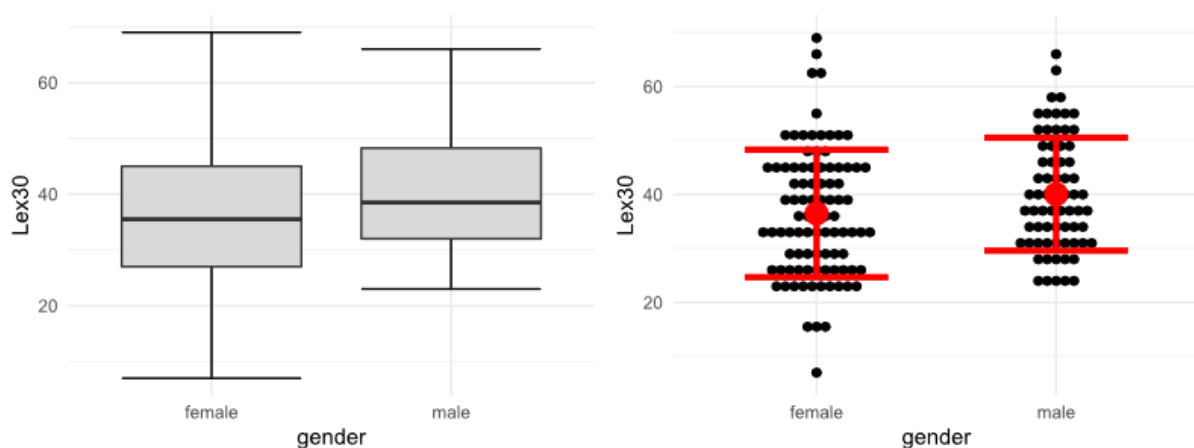
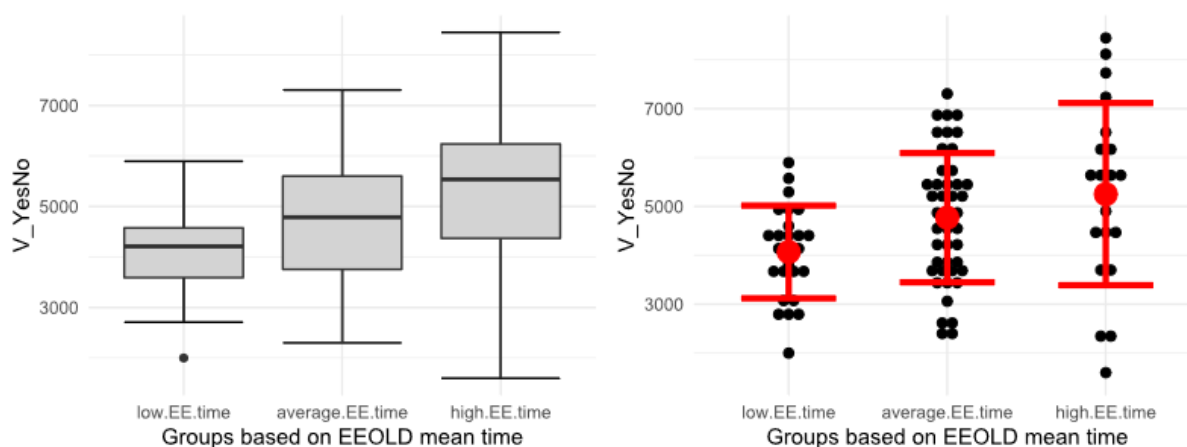*Figure 6.26: Boxplot (left) displaying median and interquartile range and dotplot (right) showing mean and standard deviation (red error bar) of Lex30 scores according to gender (N = 160)*
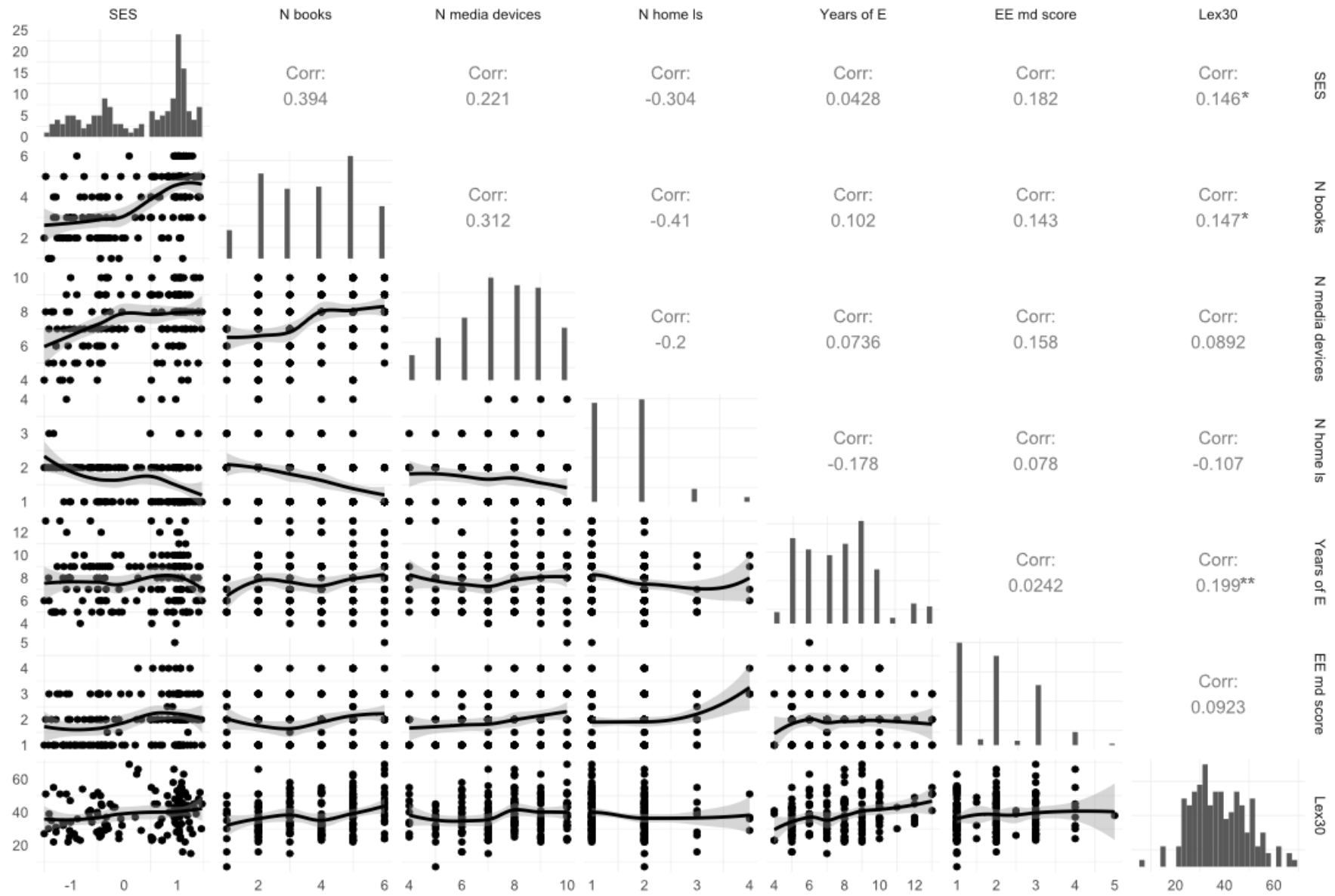
*Figure 6.27: Scatterplot matrix showing relations between influencing factors and Lex30 score graphically (lower half) and numerically through Kendall's tau (upper half, * p < .05, ** p < .01)*

Figure 6.27 presents a scatterplot matrix of the Lex30 score and six continuous variables that could potentially influence its results. As in the discussion of influencing factors and receptive vocabulary size in section 6.4.2, the main focus here is on the relationships between these participant variables and the Lex30 score since other connections have already been discussed in section 6.3.4. Beginning from the left, the SES summary variable (τ = .15 [.03, .26], p = .014) shows a very small positive and significant correlation with productive vocabulary size, as does the number of books available at students' homes (τ = .15 [.03, .26], p = .012). Access to different media (τ = .09 [-.03, .20], p = .122) is not significantly related to productive vocabulary size. Continuing with variables related to participants' language background, the plot for the number of home languages appears to indicate a slight negative relationship with productive vocabulary size (τ = -.11 [-.22, .02], p = .094), which is however not statistically significant. In contrast, length of English instruction operationalized as the number of years of English teaching (τ = .20 [.09, .29], p < .001) shows a small positive correlation with the Lex30 scores.

The primary question, however, is again whether the EE median score correlates positively with the results of the productive vocabulary test. Both the scatterplot and the coefficient in the right bottom corner suggest that there is a small correlation, but this relationship is not statistically significant (τ = .09 [-.03, .21], p = .126) and much weaker than for V_YesNo (τ = .23 [.10, .35], p < .001), see section 6.4.2). It thus appears that the frequency of EE activities shows a small, but positive and statistically significant relationship with receptive vocabulary size, but not with productive vocabulary. Another observation to be made at this point is that while the relations between the variables considered in this section are comparable to those with the V_YesNo scores described in section 6.4.2 in terms of direction, the magnitude of the correlations is generally smaller for the productive vocabulary data and some variables that show a statistically significant relationship with V_YesNo scores are not significantly related to Lex30.

Differences in Lex30 scores for groups exhibiting different EE behaviours were also examined by comparing extreme groups. Again, extreme EE groups were first created on the basis of the EE median score (see section 6.4.2). Figure 6.28 shows that the differences between the three EE groups are much smaller for Lex30 than for V_YesNo. Productive vocabulary knowledge as measured by Lex30 is greater in the high EE group (Mdn_{high} = 39 [32.5, 41.0], M_{high} = 40.12, [36.87, 43.98], SD_{high} = 11.63, n_{high} = 43) than in the low EE group (Mdn_{low} = 34 [30.91, 36.0], M_{low} =36.03, [33.26, 38.71], SD_{low} = 11.18, n_{low} = 61), but as can be seen there is hardly any difference between the average ($Mdn_{av.}$ = 39 [33.0, 44.0], $M_{av.}$ = 38.33, [35.45, 41.16], $SD_{av.}$ = 11.21, $n_{av.}$ = 57) and the high EE group. Similarly, a Kruskal-Wallis test (H = 2.54, df = 2, p = .280) shows that the differences in productive vocabulary size between the three EE groups are not statistically significant.
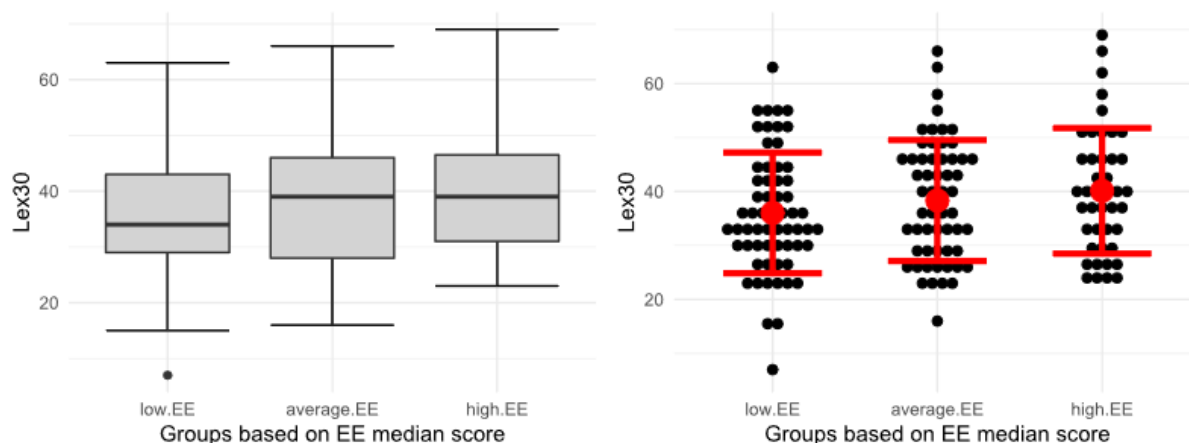
Figure 6.28: Boxplot (left) displaying median and interquartile range and dotplot (right) showing mean and standard deviation (red error bar) of Lex30 scores according to EE extreme groups based on EE median score (N = 161)

Figure 6.29 shows that when looking at the differences between groups based on time spent with EE rather than frequency of EE activities, the picture is further complicated as the average EE group has the highest Lex30 scores ($Mdn_{average}$ = 38 [33, 41], $M_{average}$ = 39.73, [36.94, 42.96], $SD_{average}$ = 10.85, $n_{average}$ = 48). The Lex30 scores are higher in the high EE group (Mdn_{high} = 37 [32, 39], M_{high} = 38.72, [34.90, 43.65], SD_{high} = 11.30, n_{high} = 25) than in the low EE group (Mdn_{low} = 32 [27, 39.5], M_{low} = 33.9, [30.52, 37.60], SD_{low} = 10.00, n_{low} = 30), but a Kruskal-Wallis test (H = 5.02, df = 2, p = 0.081) shows that the differences between the groups based on EE mean time are not statistically significant.
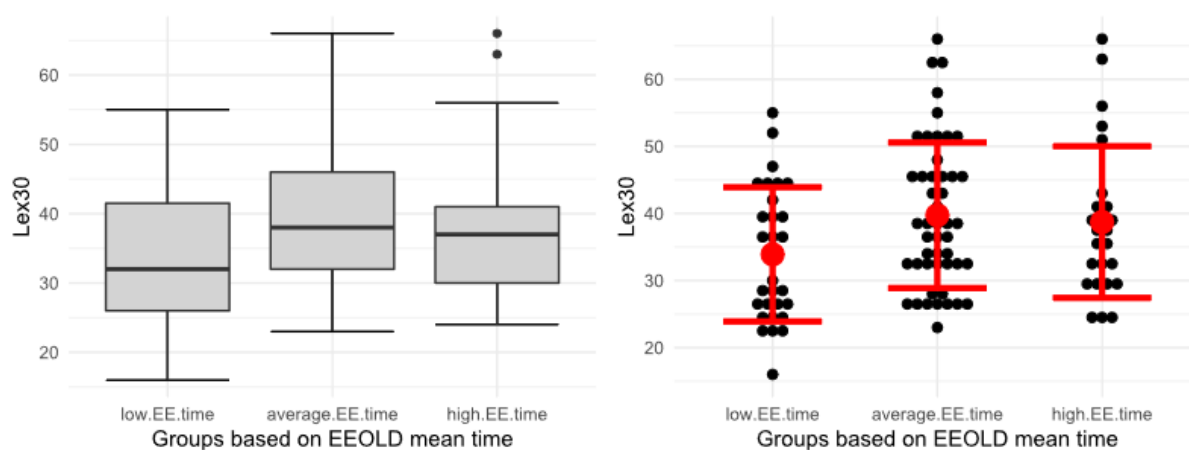
Figure 6.29: Boxplot (left) displaying median and interquartile range and dotplot (right) showing mean and standard deviation (red error bar) of Lex30 scores according to EE extreme groups based on EE mean time (N = 103)
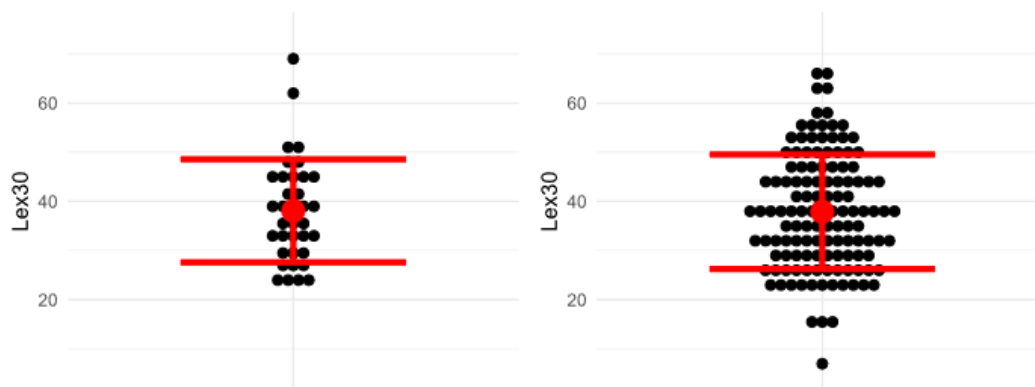
Figure 6.30: Dotplots comparing the Lex30 scores in a sub-sample of participants engaging in niche EE activities (N = 36, left) and the remaining participants (N = 125, right) showing the mean and standard deviation (red error bar)

In addition, the sub-sample of participants engaging in niche activities (see sections 6.3.1 and 6.4.2) was again compared to the overall sample. It is plausible that the 43 participants who are involved in these more demanding EE activities that require more productive language use than the most popular activities have a larger productive vocabulary size. However, when comparing descriptive statistics, we see that while the mean score in the sub-sample engaging in niche activities (M = 38.08 [35.11, 41.91], SD = 10.48, n = 36) is slightly higher than among the remaining participants (M = 37.9 [35.84, 39.98], SD = 11.64, n = 125), the median score in the sub-sample (Mdn = 37.0 [33, 41]) is exactly the same as for the remaining participants (Mdn = 37 [34, 39]). The visualization in Figure 6.30 also shows that the difference in means is negligibly small and a Wilcoxon rank-sum test (W = 2248.5, p = .503) shows that it is not statistically significant with a very small effect (r = -.05 [-.21, .10]). This analysis further supports the conclusion that extramural activities seem to play less of a role for productive vocabulary size, as niche activities, which generally involve more productive language use are not related to higher Lex30 scores either. However, methodological factors have to kept in mind as assessing productive vocabulary size is notoriously difficult (see section 3.2.3) and the test instrument used may simply not be sensitive enough.

6.4.6 Modelling productive vocabulary size

As with receptive vocabulary size, a multiple regression model with the Lex30 score as the outcome variable was computed to estimate the impact of the predictors described in section 6.4.3 (see Table 6.20). As before, computational issues emerged when attempting to build a linear mixed effects model: although the number of cases per group is slightly higher for the Lex30 data than the V_YesNo data, it was still too low to allow for valid estimation of random effects.
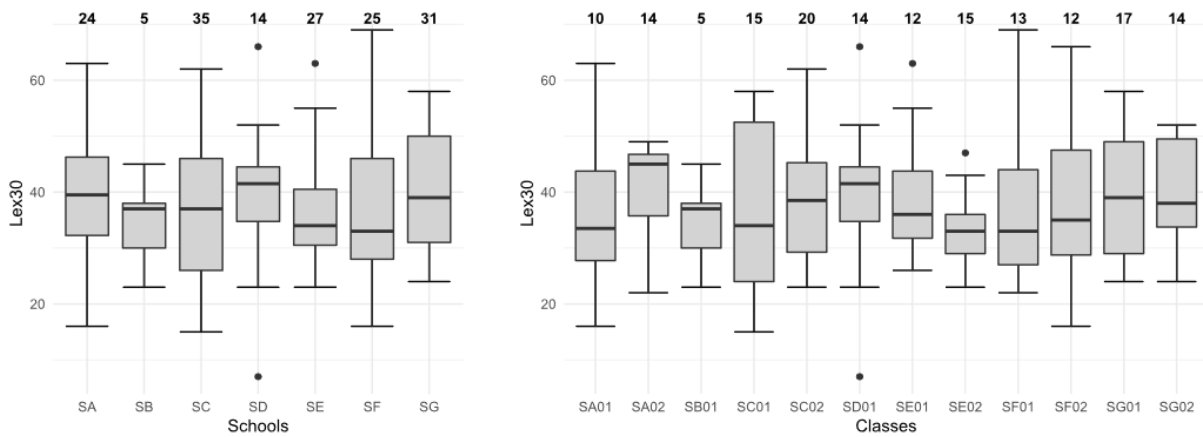
Figure 6.31: Boxplots displaying the V_YesNo scores split by the seven participating schools and twelve participating classes (N=161)

Consequently, a standard multiple regression model for productive vocabulary size was built in parallel to the models for receptive vocabulary size presented in section 6.4.3. The percentage of missing values for Lex30 as the outcome variable and the six predictor variables ranged from 0 to 17%; in total 59 out of 189 cases (31%) were incomplete. However, a complete case analysis is again preferable over multiple imputation. The model for productive vocabulary size as measured by Lex30 was therefore fit on 130 complete cases and contains the same predictors as the model for receptive vocabulary size. A summary of the model is presented in Table 6.23: it only explains 8.5% of the variance in Lex30 scores and length of English instruction is the only significant predictor if α = 0.05. Hence, in line with the bivariate analysis, frequency of engagement with EE does not emerge as a statistically significant predictor for productive vocabulary size as measured by Lex30.

	B	**95% CI**	*SE*	*p*	*β*	*sr²*
Intercept	26.42	[13.85, 38.98]	6.34	<.001**	-0.05	
EE	1.28	[-0.92, 3.48]	1.11	.251	0.10	0
Length of English instruction	1.23	[0.35, 2.11]	0.44	.006**	0.23	0.004
SES	2.03	[-0.37, 4.43]	1.21	.097	0.18	0
Media access	-0.14	[-1.34, 1.05]	0.60	.813	-0.02	0
Gender (male)	2.66	[-1.17, 6.59]	1.94	.172	0.23	0
Number of home languages	-0.05	[-3.06, 2.96]	1.52	.975	-0.00	0

N = 130, adjusted R^2 = .08 [.02, .26]

Table 6.23: Coefficients of standard multiple regression model for Lex30 scores

Again, diagnostic plots and tests were used to investigate the assumptions of multiple regression and to identify any overly influential data points. Figure 6.32 displays the continuous predictor variables in relation to the outcome variable Lex30; it illustrates that their relations can be regarded as reasonably, though not perfectly, linear. As was the case with the model for receptive vocabulary size, the assumption of independence is met as each value was produced by a different participant for both outcome and predictor variables. The Durbin-Watson test

used to test for independence of errors showed a test statistic of 2.06 (p = .766) and a value of -0.05 for autocorrelation, indicating that there is no problem (see section 6.4.3). Multicollinearity was again assessed with the help of VIF, which varies between 1.03 and 1.37 and is thus well below the critical value.
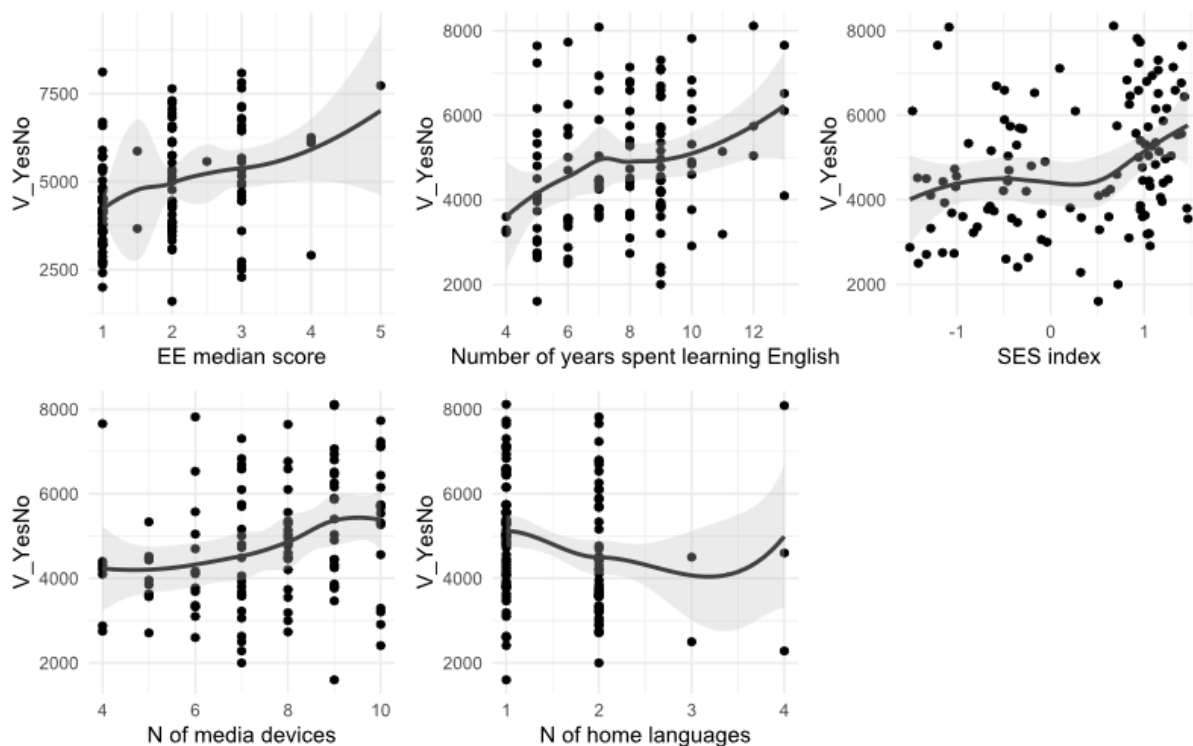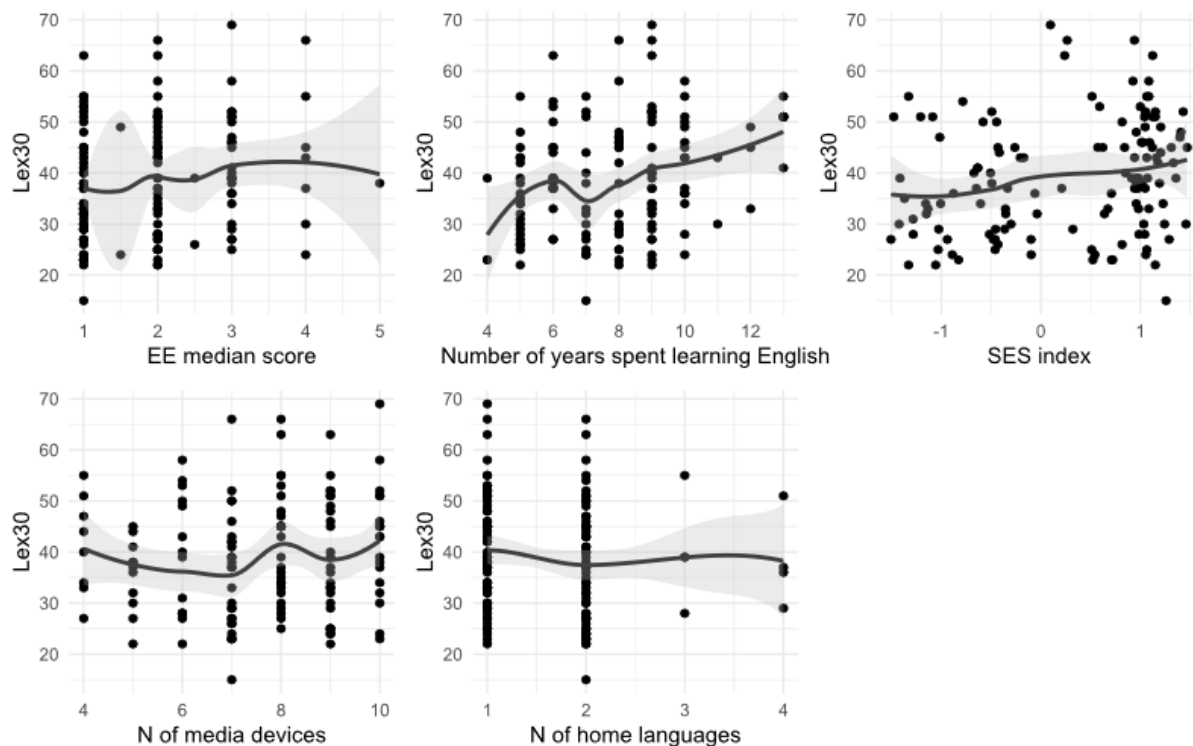

Figure 6.32: Scatterplots of the relationship between the predictors included in the regression model and the outcome variable Lex30

Assumptions concerning residuals can be explored with the help of Figure 6.33. The first plot of residuals against fitted values shows that the assumption of homogeneity of variance has been met, while the normal Q-Q plot indicates that the residuals are fairly close to a normal distribution. The residual vs. leverage plot can be used to detect outliers and data points that exert an undue amount of influence on the model, but since the dotted line for Cook's distance is not visible such concerns are unnecessary. Indeed, the maximum of Cook's distance displayed in the fourth plot is below 0.07 and thus far below the critical value of 1.

The multiple regression analysis for productive vocabulary size shows that only length of English instruction is a significant predictor of the Lex30 scores, but the model explains little variation in scores overall. In combination with the results on receptive vocabulary size presented in section 6.4.3, this finding indicates that extramural English has a much larger effect on receptive than productive vocabulary size and that the number of years spent learning English is the only variable that significantly predicts both receptive and productive vocabulary size.
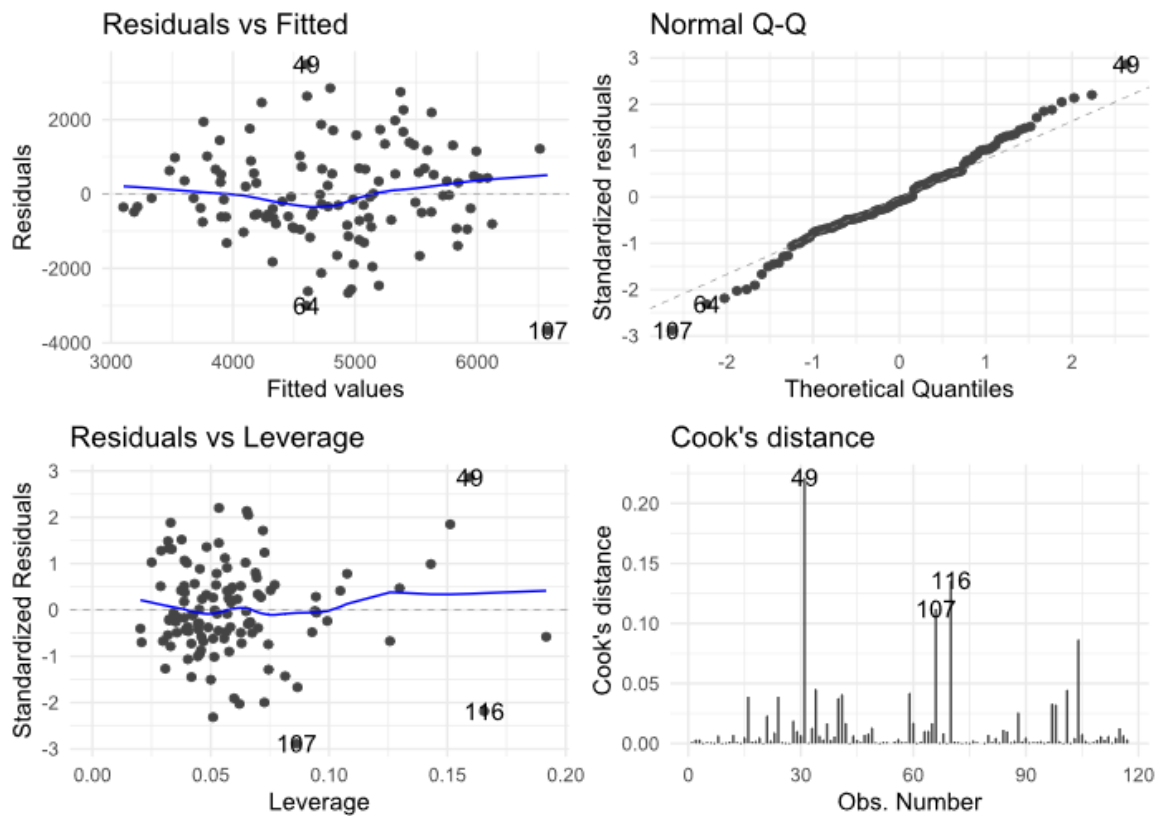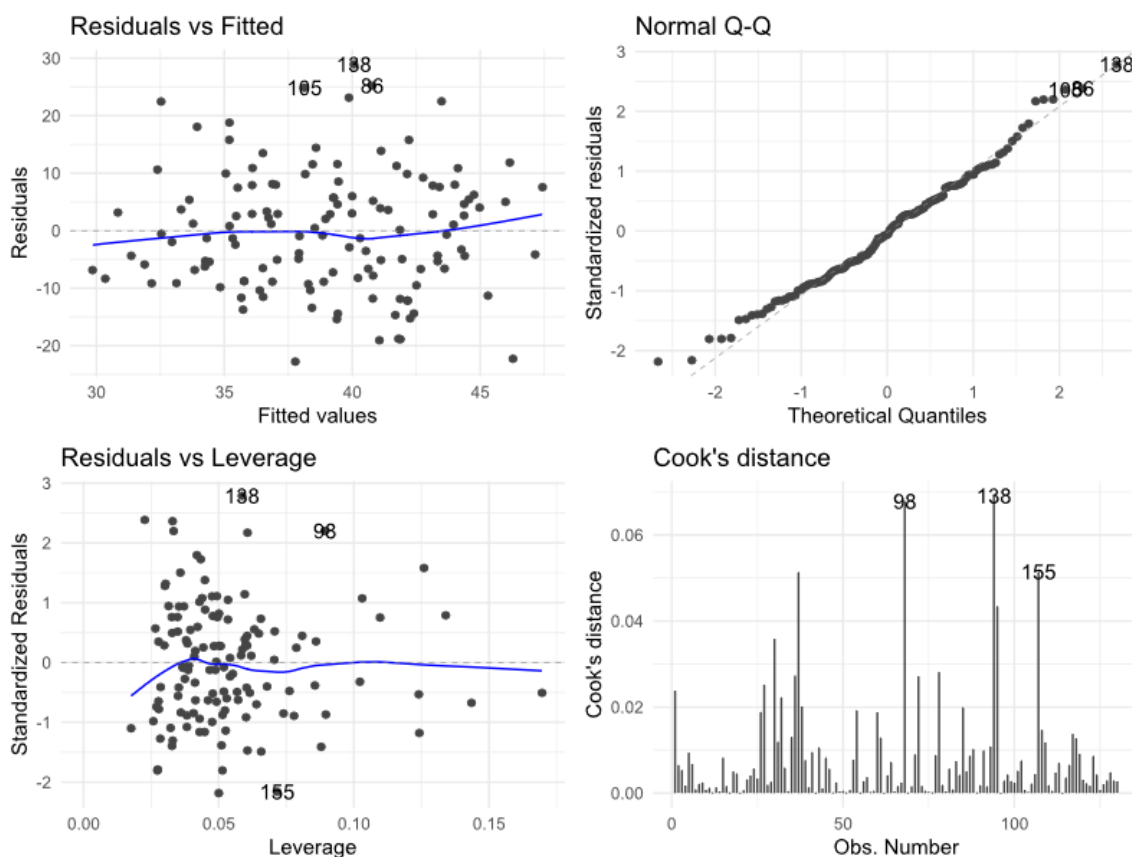
Figure 6.33: Diagnostic plots for the linear regression model: residuals vs fitted values (upper left), normal Q-Q plot of residuals (upper right), residuals vs leverage plot (lower left) and Cook's distance (lower left)

However, although the results of both the bivariate and multivariate analyses indicate that in contrast to receptive vocabulary size there is no relationship between productive vocabulary size and extramural English, an additional exploratory analysis of the Lex30 data presented in the next section seems to suggest that exposure to EE does have some effect on productive vocabulary knowledge as well.

6.4.7 Exploring connections between productive vocabulary and extramural English further: the schoolbook analysis

In their 2010 article Fitzpatrick and Clenton made the following statement in an article on the reliability and validity of Lex30:

> [… A]s a research tool we would argue that the test has great potential on two counts. First, it can be used in its standard and intended form as, essentially, a frequency-based vocabulary measure. Second, the word association task component can stand alone as an effective and efficient elicitation tool, which can be combined with a range of analytical measures (Fitzpatrick & Clenton 2010: 551).

So far, this possibility of combining lexical samples elicited with the help of the Lex30 word association task with analytical measures other than the standard frequency-based scoring procedure has, to the best of my knowledge, not been taken up in research studies. This section describes a first endeavour to put Fitzpatrick and Clenton's (2010) suggestion into practice by comparing samples elicited with the help of Lex30 to vocabulary presented in the participants' schoolbooks and thereby attempting to draw conclusions about nature of vocabulary likely to

have been acquired outside the classroom. The questions investigated in this additional analysis thus are

a) *How many response words produced in the Lex30 test were not included in participants' schoolbooks?*

b) *What, if anything, can these additional words tell us about teenagers' productive vocabulary knowledge and its relation to extramural English?*

It has to be emphasized that this analysis is of a highly exploratory nature and that, as any first attempt to venture into new methodological territory, it is subject to methodological limitations. The most apparent limitation is that extramural here is operationalized as 'extra-coursebook', despite the fact that it is commonly the case, and definitely desirable, that more lexical items than those presented in a coursebook are taught by English teachers in their classrooms. It was however practically impossible to collect further data on vocabulary taught, like for instance vocabulary notebooks, in the 12 participating classes as these would have needed to be collected from their first year in secondary school (grade 5) onwards. Indeed, this exploratory analysis is only possible because solely schools which offer the lower and upper levels of secondary education from grade 5 to 12 were selected for this study (see sections 4.2.1 and 5.3.1). It can thus be assumed that all participants in one class were taught using the same lower and upper secondary school coursebooks, unless they joined the school later, which is however relatively unlikely because schools offering both the lower and upper level rarely accept new students after grade 5.[183]

The participating classes' current English teachers provided information on the lower secondary coursebook their students had used in grades 5 to 8, the coursebook they were using at upper secondary level, and on which units they had already covered in grade 10 at the time of data collection. In total, the 12 classes had been taught using two coursebooks for lower secondary school, the very popular *More* series (Gerngross et al. 2008a, 2008b, 2009a, 2009b) and *Red Line* (Haß 2007, 2008a, 2008b, 2009), and three different schoolbooks for upper secondary level, *English in Context* (Abram & Williams 2016, 2017), *Into English* (Puchta et al. 2013a, 2013b) and *Prime Time* (Hellmayr, Waba & Mlakar 2010a, 2010b). Consequently, eight volumes of lower secondary cousebooks (grade 5 to 8) and six volumes of upper secondary coursebooks (grade 9 and 10) needed to be taken into account in the analysis. Lists of lexical items presented in these coursebooks had to be prepared so that they could be used as reference lists against which to compare the Lex30 samples. Compiling these vocabulary reference lists from the schoolbooks turned out to be a rather lengthy process: first, the vocabulary presented at the end of every unit of a given volume (e.g. 'word field' and 'words and phrases') was entered into an Excel file and then cleaned and lemmatized:[184] cleaning was made necessary as lexical items were presented

---

[183] In addition, three participants who had only recently moved to Austria and of two more students who had attended a different school providing EMI education in lower secondary were excluded from analyses (see section 5.3.2).

[184] Some of the word lists could be found as digital versions either on companion websites or on websites where teachers had produced quizzes for their classes and only had to be checked manually, whereas the remainder had to be typed into Excel.

repeatedly in different units and lemmatizing was essential as the Lex30 samples also had to be lemmatized for the regular frequency-based analysis. In order to ensure comparability, the same criteria used for the Lex30 samples were also applied to the lemmatization of the schoolbook lists and to the treatment of multiword units, which were split up into their constituent parts (see section 5.3.5.4). Once the first volume of a coursebook series was completed, the same process was carried out for all other volumes and these lists were then checked for repetitions as well.[185] For the two lower secondary coursebooks, it was found that the unit vocabulary frequently did not include very prominent words presented, which was why all words presented in the 'wordlist' or 'dictionary', a sort of glossary for learners, at the back of the coursebooks were compared to the unit vocabulary lists and those items not included in the latter were added as well.

Once the reference lists had been established from the five coursebook series, AntWordProfiler (version 1.4.0w) was used to compare them with the 172 Lex30 samples collected in this study (see sections 5.3.5.4 and 6.4.4). For each class, the samples were first compared to the reference list containing the vocabulary presented in the lower secondary coursebook they had used. The offlist types, i.e. those types produced by participants that were not included in the lower secondary reference lists, were then recorded in a new file and run against the upper secondary level list for this class. The reference list based on the upper secondary coursebook used in a given class always contained the entire vocabulary of the first volume (grade 9) and only those units of the second volume (grade 10) that the group had been taught in class before the administration of the vocabulary tests. This customization of the reference lists based on the individual progress of classes was carried out to account for differences between the groups of learners so that additional lexical items not yet encountered in the schoolbook could be filtered out as accurately as possible. Following the same procedural steps, it is not only possible to produce lists of offlist types for whole classes, but also for individual students. In principle then, the analysis presented here could be taken further to compare offlist types produced by groups of students from different classes who engage in similar EE behaviours to explore any potential commonalities; however, for reasons of time and practicality such an analysis was not carried out in the present study.

Turning back to the two questions posed at the beginning of this section, the results of the schoolbook analysis show that the 172 participants whose samples were analysed produced 15,660 tokens and 2,207 types overall. Of these, 1,469 tokens and 801 types were not included in their respective schoolbooks. [186] Based on the separate analyses for the 12 school classes, which

---

[185] On a methodological note, AntWordProfiler also proved very useful in compiling these lists: the programme does not allow the reference lists ('level lists') to contain duplicates; hence, loading several lists to be compared as level lists results in the programme flagging any recurring items. In addition, several lists can easily be combined by loading them as user files and running them against an empty dummy level list, in this case the programme produces a list of all types in alphabetical order when the option "include words in user file(s) but not in level list(s)" is selected.

[186] This total number of types potentially contains lexical items that were presented in some of the students' schoolbooks but not in others because the number results from a combination of the offlist types based on the analyses per class.

are summarized in Table 6.24, the average number of types included in the schoolbooks was 80.02%, which means that 19.98% (± 3.18%) of the types produced by participants on the Lex30 test were not included in their respective schoolbooks. As visualized in Figure 6.34, almost one fifth of the response words were not found in students' coursebooks indicating that it is very likely that many of these types were acquired in out-of-school contexts, even if some of them may have been explicitly taught by teachers.

	N of students	Total number of tokens/types produced		Number of tokens/types not included in schoolbooks		Proportion of types not included
		tokens	types	tokens	types	
SA01	10	921	488	131	92	18.85%
SA02	15	1381	641	214	143	22.31%
SB01	7	668	398	87	74	18.59%
SC01	17	1534	698	203	145	20.77%
SC02	21	1869	751	277	170	22.64%
SD01	15	1387	614	173	132	21.50%
SE01	13	1193	639	166	130	20.34%
SE02	15	1420	601	148	101	16.81%
SF01	13	1134	574	148	114	19.86%
SF02	13	1136	598	143	115	19.23%
SG01	18	1601	651	190	136	20.89%
SG02	15	1516	650	164	117	18.00%

Table 6.24: Descriptive results of Lex30 schoolbook analysis per class: Number of students, total number of types and tokens produced, number of types and tokens not included in the respective schoolbooks, proportion of types not included in the schoolbooks
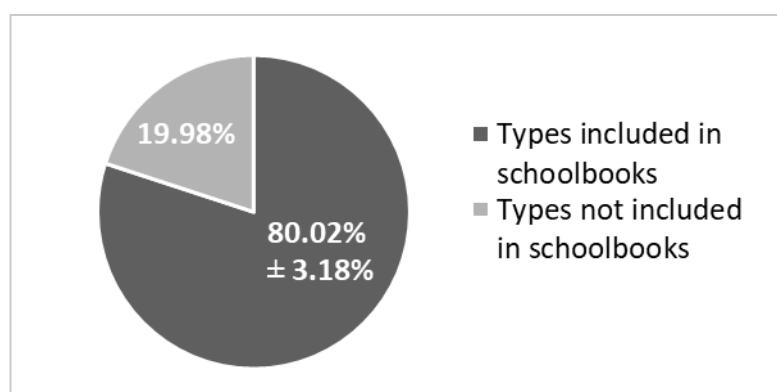
Figure 6.34: Proportion of types produced on the Lex30 test (not) included in participants' coursebooks (average across all classes ± 3.18%)

The second question relates to the kinds of lexical items that students produced although they were not presented in their schoolbooks, and possible connections between these lexical items and EE activities. A frequency analysis using Nation's (2012a) BNC/COCA lists reveals that a large majority of the types identified in the schoolbook analysis fall into the category of mid-frequency vocabulary as shown in Figure 6.35. According to Nation's (2013: 16–19) division into high-, mid- and low-frequency vocabulary (see also Schmitt & Schmitt 2014) approximately 27% of the types

are among the 2,000 most common English word families and thus highly frequent, just over 60% are in the range of the 3000 to 9000 frequency bands, and about 6% each are either very infrequent or were not found in the BNC/COCA reference lists.
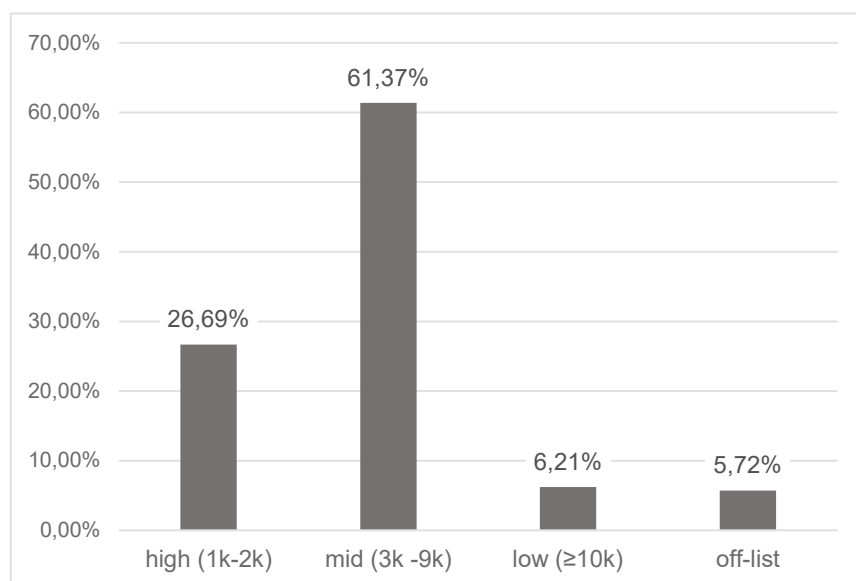
Figure 6.35: Frequency of offlist types identified in schoolbook analysis using Nation's (2012) 25k BNC/COCA lists

The relatively large amount of high frequency words that is not included in students' coursebooks is surprising as one would expect that schoolbooks guiding almost six years of English teaching would present these common and extremely useful words first. Examples of word types produced by the participants, but not found in at least one of the coursebooks included *brain, calculate, dirt, glasses, selfish, payment* or *ugly*, which are arguably a valuable resource for understanding and taking part in everyday discourse.[187] This finding thus suggests either that the writers of these schoolbooks heavily rely on incidental learning of unpresented high-frequency vocabulary or that the design of the coursebooks does not follow a principled approach when it comes to the presentation of high-frequency vocabulary, a situation that has unfortunately also been found in other studies (e.g. O'Loughlin 2012).

The vast majority of types produced in the Lex30 test that are not included in participants' schoolbooks are part of a mid-frequency band of vocabulary, as mentioned above. These words are "needed to deal with English without the need for outside support" (Nation 2013: 18). According to coverage research discussed in section 3.1.2, a minimum of 4,000 word families are needed to read unsimplified texts and at least 3,000 word families are necessary to understand spoken English on TV. Hence, these words are firstly very useful for learners at an (upper-)intermediate level such as the participants in the present study and secondly, they are essential for engaging in extramural English activities. Additionally, it is plausible that participants

---

[187] Being able to communicate in the L2 in everyday life situations in manner appropriate to age and level of learning is an aim specified in the curriculum for all foreign languages in the lower secondary AHS curriculum (Bundesministerium für Unterricht und kulturelle Angelegenheiten 2000).

frequently encounter such mid-frequency words in English-language input in out-of-school contexts and are thus able to acquire some aspects of word knowledge for these.

An analysis of the parts of speech among types not found in schoolbooks revealed that most of them are nouns: 61.45% are nouns (e.g. *atom, exhaustion, weapon*); in addition, 10.81% of words can be used as a noun or a verb (e.g. *mail, stalk, trash*) and 2.29% can be used as a noun or adjective (e.g. *antibiotic, explosive, original*). In contrast, only 10.94% of the offlist types are verbs (e.g. *browse, complicate, enchant*), 13.99% are adjectives (*elementary, pure, wise*) and 0.25% are adverbs (*aboard, quite*).[188] The preference for nouns could stem from two aspects: test characteristics and intrinsic word difficulty. While the word association task may tend to elicit more nouns than other part of speech classes, nouns have also been shown to be more easily acquired (see section 3.1.3).

A third aspect of analysis is the categorization of words produced on the Lex30 test but not found in students' coursebooks in terms of thematic fields, which might point to possible connections with EE. All offlist types were coded for topic areas manually by the researcher in a bottom-up process that led to the establishment of 23 thematic categories. Due to practical constraints it was not possible to engage a second rater, but coding and reviewing previous codes at several points of time ensured that all judgements were subject to repeated scrutiny. The tentative nature of this analysis has to be highlighted again because this first attempt at a thematic analysis of offlist types cannot lead to any definite results. Nonetheless, in the spirit of exploring the Lex30 data in as much depth as possible the thematic analysis can potentially point to aspects worth further investigation.

Table 6.25 presents the 23 thematic fields identified; in total 450 types were assigned to categories in this analysis. In 27 cases in which a word could be regarded as part of two topic areas a secondary category was assigned as well; thus, there is some overlap between the thematic fields (for a complete list of all types assigned to the thematic fields please see Table B.15 in Appendix B). The largest thematic field found among the 801 types was *food* with 57 items followed by *medicine*; *beauty & clothes; crime, terror & war*; *science* and *feelings.* Clearly, some of the thematic areas identified are conditioned by the cue words on the Lex30 test: for instance, *cloth* elicited many answers relating to clothes, *disease* those relating to medicine, *fruit, potato* and *rice* often led to food being named and *furniture* or *science* also frequently elicited topical vocabulary.

---

[188] *aboard* can also be used as a preposition, but since it was the only preposition found, no extra category was established.

Thematic fields	As primary category	As secondary category	Total number of word types assigned
Food	56	1	57
Medicine	35	4	39
Beauty & clothes	34	0	34
Crime, terror & war	30	1	31
Science	28	2	30
Feelings	29	1	30
Character	27	1	28
Body	23	1	24
Death	17	3	20
Politics	18	1	19
Economy	17	2	19
Fantasy	17	1	18
Historical warfare	14	3	17
Sports	16	0	16
Furniture	14	0	14
Pejorative terms	11	2	13
Religion & beliefs	12	0	12
Nature	12	0	12
Media	9	2	11
Computer & technology	9	1	10
Work & jobs	9	0	9
Animals	8	0	8
Drugs	5	1	6

Table 6.25: Thematic fields identified among the types not found in participants' schoolbooks

Other thematic areas are however not easily explained by a topical bias due to cue words, for instance the three thematic fields relating to violence:[189] *crime, terror & war*, which comprises terms relating to criminal behaviour (e.g. *mafia, scam, stalk*) as well as modern warfare and terrorism (e.g. *assault, combat, explosive, loot, sniper*); *historical warfare*, which includes vocabulary related to ancient and more primitive warfare (e.g. *conquer, foe, fortress, saber, siege, sword, warrior*) and *death*, which relates to ways of dying (e.g. *choke, drown, mutilate, starvation, strangle*) and concepts related to death (e.g. *cemetery, coffin, corpse, graveyard, grief, immortal*). This abundance of vocabulary relating to war and death is surprising and could point to students' extramural environments as a source for vocabulary learning: many films and mainstream TV series deal with topics of crime, terrorism and espionage, first-person shooter and strategy games are hugely popular, and daily news reports also present and recycle much of this vocabulary. Especially in the case of vocabulary relating to historical warfare it seems reasonable that participants picked up many of the words from novels, games or TV series dealing with historical or fantastical wars such as the currently hugely popular series *Game of Thrones*. This thematic field might thus be connected to *fantasy,* another topic area which can

---

[189] The only cue word which frequently elicited vocabulary related to violence was *attack,* but interestingly response words in the three word fields described were also produced in response to other cue words.

plausibly be connected to out-of-school entertainment and which includes words such as *alchemy, enchant, potion, sorcery, wand, warlock* or *witchcraft*. One last thematic field that is most probably related to out-of-school contexts as these words would not be taught in English classes are pejorative terms like *creep, dick, idiot, moron, retard* or *snitch*.

While the results of the thematic analysis are certainly interesting, any connections to extramural English remain hypothetical at this point, although it would certainly be worthwhile to further investigate what kind of vocabulary teenagers learn outside the classroom, whether incidentally or intentionally (see also Chapter 7). One last aspect that highlights a connection between the Lex30 samples and the wider world participants live in are explicit references to pop culture, including brands and titles found in the data. According to the Lex30 scoring criteria these were excluded from analysis as proper nouns, but it is nonetheless of interest that they came to participants' minds while having to produce associations in a limited amount of time. The most frequently mentioned proper noun was *Google* as a response to *map* (14 times), but *Dora* as in *Dora the Explorer* from the eponymous animated series was also named six times.[190] The cue word *obey* resulted in response words like *brand, cap* or *sweatshirt* 13 times, which clearly points to the clothing brand *Obey* rather than any semantic meaning of the word.[191] *Harry Potter* was given as a response to *spell* 11 times, *dirty* elicited *dancing* as in the film title six times, *real* made participants think of the football team *Real Madrid* six times and *television* evoked *Netflix* four times.[192] These examples show, perhaps not unexpectedly, that titles and brand names form strong association in participants' minds, but in fact the association could also be cultural: the cue word *potato* led to five responses referring to *9Gag*, an internet portal showing funny memes and video clips. It appears that this association stems from a practice on the 9Gag site to post a picture of a potato underneath especially long posts to reward the reader for making it to the end, thus, an internet culture phenomenon also influenced participants' responses in the Lex30 association task.[193] Taken together, these examples show that test takers' living environment has an impact on test performance, and while this is probably true for any test, it is especially salient in Lex30 due to the word association task used for the elicitation of language samples.

In summary, on average approximately one fifth of the types produced in the Lex30 task were not included in participants' schoolbooks. Most of these words are nouns and fall into the category of mid-frequency vocabulary; thus, these are words that are used relatively frequently in authentic language use. This result supports the hypothesis that many of the offlist types identified in the schoolbook analysis have been acquired through English-language input in

---

[190] A map plays an important role in the *Dora the Explorer* series and even has its own song.

[191] Obey is a clothing brand founded by Shepard Fairey in 2001, which sees itself routed in counterculture, for further information please visit https://obeyclothing.eu/.

[192] From a methodological perspective it is interesting to note here that while the EEQ differentiated between watching series on TV and online, *Netflix* as a response to the cue word *television* indicates that such a distinction between TV and online streaming does not necessarily exist in the minds of young people and that television is probably increasingly used to refer to any form of audiovisual broadcast.

[193] see for instance https://acumagnet.wordpress.com/2014/12/29/sorry-for-the-long-post-heres-a-potato-comes-from-glados-in-portal-2/

extramural contexts. A thematic analysis of the types not included in coursebooks shows that evidently cue words predetermine topic areas to some extent, but thematic fields related to death, modern and historical warfare, fantasy and pejorative terms suggest that extramural input also plays a role. Overall, this exploratory analysis produced interesting, but highly tentative results and indicates that further research into the kinds of (productive) vocabulary learned outside school would be a worthwhile endeavour.

6.5   Summary

This chapter has presented the results of the quantitative strand based on the analysis of the combined data from the EEQ, the EEOLD, Lex30 and V_YesNo. In terms of background variables, the questionnaire data show that the sample of Viennese 10th-grade students participating in this study are comparable to the wider student population in terms of gender and age, that more than half are multilingual, and that the socioeconomic status of their families is above average in comparison to Austrian data collected in large-scale studies. Concerning English, the vast majority began studying it at school, as is typical for the Austrian context, and over 90% now rate their English proficiency as CEFR level B1 or higher on the self-assessment scale. Participants generally have positive attitudes towards English and report encountering it most frequently on the internet and on social media, as well as in films, series, and video clips; in fact, over 65% state that they use it more during their leisure time than in school lessons.

Data on engagement with extramural English shows that over 95% come in contact with English almost on a daily basis and that teenagers' preferred EE activities coincide with their generally preferred leisure activities and are subject to similar gender differences. Music, audiovisual (online) media, and other online contexts are the most popular activities, but the range of activities participants engage in is highly individualized and impressively varied and appears to be determined by specialized interests and personal preferences to a large extent. The most common activities are mainly carried out in online contexts and mostly involve receptive language use, while more infrequent activities and specialized niche activities often also entail language production. With regard to time spent with English outside school, the EEOLD data suggest that students spend a large part of their leisure time with English with a mean of approximately four hours per day. This remarkable finding may seem surprising but is supported by further analyses with regard to time use on weekends and the role of music, as well as qualitative data from the focus group interviews (see section 7.2.2). Although engagement with EE clearly is the result of several factors ranging from global to personal, aesthetic reasons and the desire to read or watch original versions seem to be the most common motives for students' use of English in their free time.

Regarding receptive vocabulary knowledge, a number of different analyses of the V_YesNo data were carried out. The results indicate that Viennese 10th-grade students in academic secondary schools have a mean vocabulary size between 3,900 ($h×CJ\%$ score) and 4,800 lemmas (V_YesNo score) depending on the scoring method. Male participants have larger receptive vocabulary

sizes than female ones, but the difference is not statistically significant. Frequency of engagement with extramural English, length of instruction and all three SES-related variables show significant positive relationships with the receptive test scores. In a multiple regression model including six variables EE and length of instruction emerged as significant predictors of receptive vocabulary size with frequency of engagement with EE explaining slightly more variance in test scores than length of instruction.

Productive vocabulary size as measured by Lex30 shows significant positive correlations with the SES summary index, the number of books available at home and length of English instruction. In addition, boys again outperformed girls with the difference being statistically significant for Lex30 scores. However, there is no statistically significant correlation between Lex30 and other variables including the EE median score. Similarly, length of instruction is the only significant predictor in the multiple regression model for productive vocabulary size. The data therefore suggest that frequency of engagement with English in informal contexts outside school is more clearly related to receptive than productive vocabulary knowledge. However, an exploratory analysis, which compares data generated with the help of Lex30 against the vocabulary presented in students' schoolbooks, shows that about 20% of the response words produced in the Lex30 task did not come up in participants' coursebooks. Further thematic analysis indicates that it is likely that at least some of these words have been learned in out-of-of school contexts. Together with the results of the regression modelling this finding could be taken to mean that there also is some effect of extramural English on productive word knowledge, but that it is quantitatively and qualitatively different from the effect on receptive vocabulary knowledge, for which the relationship appears to be much stronger.

7   Results of the qualitative strand

This chapter presents results relating to the overall qualitative research question:

> **RQ 5:** *What are Viennese upper secondary school students' perceptions of EE and its potential for language learning?*

First, participants' descriptions of the importance of English in their everyday lives are summarized (section 7.1) before turning to accounts of their EE practices and evaluations of selected quantitative findings on extramural English (section 7.2). Participants' views on learning from extramural English constituted the second content focus in the group interviews (section 7.3) and in line with the aims of this study their practices with regard to vocabulary in out-of-school contexts are of particular interest (section 7.4). Finally, participants' thoughts on the relationship between in- and out-of-school English practices (section 7.5) provide interesting insights into learners' perceptions of different learning environments and the (lack of) links between these.

In the following, the contents of the focus group interviews are summarized in relation to the sub-questions for the qualitative strand specified in section 5.4. The presentation of findings again follows recommendations by Kuckartz (2016: 218–222) and uses functions provided in the analysis software MAXQDA (VERBI Software 2017) to visualize the data; for instance, each of the five sections begins with a word cloud displaying the hundred most frequent types in the main category described to immediately render the most important themes visible.[194] Since the interviews were conducted in German (see section 5.4.2) participants' comments had to be transferred into English: in the running text statements are summarized in English and cited indirectly, whereas longer quotes are provided in the original German version followed by an English translation.

7.1   The significance of English in participants' everyday lives

The answer to the first sub-question *How do participants describe the importance of English in their everyday life?* (RQ 5a) entails two parts: the evaluation of the role of English in young Austrians' everyday life as well as reasons for its importance. The related aspect of comparisons of English to other languages is discussed in section 7.1.1.

Not surprisingly, all participants evaluate English as important or even very important for young Austrians, although their evaluations of the degree of significance vary. Some participants see English as having become 'normal' in Austria: it is described, for instance, as a fixed part (Elisa, SG02: 6)[195] or as being so ordinary that it is sometimes even used unconsciously (Paul SC01: 928). Others evaluate English as not being necessary for everyday life in Austria and state that not knowing English would not be a major obstacle but, except for Franz (SE01: 32-34), who

---

[194] Since the word cloud function in MAXQDA does not currently allow individual formatting, the data were exported and the word clouds included in this chapter were produced using the website www.wortwolken.com based on MAXQDA data.

[195] References to the interview data consist of the speaker's self-chosen pseudonym, the focus group and the paragraph number, which usually corresponds to a speaker's turn.

reports not using English much in her free time, they all qualify such statements saying that while English is not strictly necessary, it is still useful or makes things easier (Walküre SF01: 70) and that they would miss something if they didn't know the language (Franz M. SC01: 45).[196]

Figure 7.1: Word cloud of the 100 most frequent types in the main category SIGNIFICANCE OF ENGLISH IN EVERYDAY LIFE OF YOUNG AUSTRIANS

Some participants even state plainly that not knowing English would be a horror (Kirito, SE01: 68), very difficult and almost impossible (KingKong & DJ SF01:51:54) or at the very least greatly limiting (Paul SC01: 46-49), especially in online contexts. This stance is most dramatically expressed by Elisa, who makes the following statement while talking about the interviewer's question to imagine what it would be like to be an Austrian teenager who did not know English:

> **Elisa**: Das war so wie eine Frage ‚Stell, stell, stell dir vor du hast einen Arm nicht. Was ändert sich für dich?
> **Interviewer**: @
> **Elisa**: Also es ist (.) das kann man nicht wegkriegen.
> **SX-m**: (Das ist halt) ein Teil vom Leben. (SG02: 51:54)
>
> **Elisa**: That is a question like 'Imagine, imagine you do not have an arm. What changes for you?'
> **Interviewer**: @
> **Elisa**: Well it is (.) you can't get rid of it.
> **SX-m**: (That's just) a part of life[197]

In relation to their statements on the significance of English participants also give reasons why English is important in the lifeworlds of young Austrians. As can be seen from Table 7.1, the reasons are manifold and concern several different aspects, such as the role of English as a lingua franca for international communication, its importance in the job world, its influence on German and its use for media consumption.

---

[196] Please note that since no restrictions were imposed on the participants' choice of pseudonyms, some of them chose names which disguise their gender or ethnic background, as is for instance the case with Franz. For information on the participants please see section 5.4.4.

[197] For an overview of the symbols used, please see the transcription conventions in Appendix A.

Subcategory	Number of coded segments [198]	Number of participants
Needed for leisure activities	21	17
It's the universal language (*die Universalsprache*)	15	13
Finding information/staying up to date	6	6
For stays abroad/holidays	5	4
Importance for the future	4	4
(International) communication	4	4
Anglicisms in German	3	3
Part of youth language	2	2

Table 7.1: Frequencies of occurrence for the subcategories of REASONS WHY ENGLISH IS IMPORTANT FOR YOUNG AUSTRIANS

In five out of six focus groups English is portrayed as the international lingua franca which is spoken everywhere and which everyone has to know nowadays, a role aptly described by Maria as *"die Universalsprache"* [the universal language] (SC01:942). This function of English is frequently mentioned in relation to online environments or international contexts, where English has to be used to make oneself understood according to participants. However, the following example illustrates that the role of English as a 'universal language' also has an impact on participants' daily life in Austria:

> **Interviewer**: [...] wie wichtig ist Englisch in eurem Alltagsleben im Vergleich zu anderen Sprachen? (2)
> **Keanu**: Ist die Weltsprache.
> **SS**: @@@
> **Keanu**: Ja ist ja so. Die Sprache, ich (2) die muss man halt können.
> **Interviewer**: In deinem Alltag ist es wichtig, weil es die Weltsprache ist?
> **Keanu**: Ja. [...]
> **Keanu**: Wenn ich auf der Straße geh und da kommen irgendwelche Ausländer und dann fragen sie mich was, und (.) dann ist das halt so, dass sie mich nicht auf Chinesisch was fragen, sondern auf Englisch. (SA02: 1-10)

> **Interviewer**: [...] how important is English in your everyday life in comparison to other languages?
> **Keanu**: It's the world language.
> **SS**: @@@
> **Keanu**: Yes, it is just like that. The language, I (2) you just have to know it.
> **Interviewer**: In your everyday life it's important because it is the world language?
> **Keanu**: Yes. [...]
> **Keanu**: If I walk on the street and then a foreigner comes and asks me something and (.) then it's not like they ask me in Chinese, but in English.

Keanu's example shows that English is regarded as the default language for communication with people who do not speak German and as such English as a lingua franca also plays a role in everyday circumstances.

Even more often participants refer to English as being needed for their leisure time activities. While these activities and the reasons why English is used for them are discussed in more detail

---

[198] The two counts included in this table as well as subsequent ones in this chapter represent the number of segments in the focus group that were coded with a specific subcategory and the number of participants who made a statement included in this category. Since one coded segment can include several speakers or speakers can comment on the topic of a subcategory more than once these two counts represent different information and do not necessarily match.

in section 7.2, it is important to point out here that EE activities are one of the aspects that increase the significance of English in young Austrians' lives. Many participants mention their spare time activities when asked to describe the importance of English in their daily lives. More specifically, they regard it as necessary for common activities such as listening to music (John SE01: 28-29) or being able to watch and read original versions (e.g. Paul SC01: 4-6, Vanessa SD01: 22, Kirito SE01: 3, Johannes SG02: 20), for their online activities and social media (e.g. Elisa SG02: 8), and for a few very specialized activities such as programming or playing basketball (e.g. Mito SD01: 14-16, DJ SF01: 2-4, see section 7.2.1). Related to leisure activities is the argument that participants need English to find information and to stay informed. As Kirito (SE01: 78) puts it, the main drawback of not knowing English would be not being up to date anymore. This reason is mentioned mostly in connection to TV series and online media (Karl SE01: 31), but also in relation to finding information with the help of search engines such as Google because more information is available in English than in German (Mito SD01: 71-78, see also section 7.2.3).

In addition, participants report needing English for communication with international friends (Anna SD01: 25-29) or friends who do not know German well (KingKong SF01: 66-68). A further argument that is mentioned in several of the groups, though not directly related to their everyday lives within Austria, is using English while on holiday or during other stays abroad (e.g. Lia SA02: 17, Emma SD01: 6, Johannes SG02: 20). In one focus group (SE01) the importance of English for the participants' future lives and in particular for their future careers comes up several times.

> **Karl**: [...] ich finde Englisch jetzt auch nicht nur auch wichtig wegen den Serien, sondern auch weil Englisch auch eine all- allgemeine Sprache ist, in jedem Land gesprochen wird und ich find dass das dann auch später auf jeden Fall helfen wird, wenn man halt Englisch gut kann und ausgebildet ist. (SE01: 31)
>
> **Karl**: [...] I do not think that English is just important because of series, but also because English is a gen- general language, it is spoken in every country and I think that it will also definitely help later, if you speak English well and are educated.

Karl's quote exemplifies an agreement in this group that English is a necessary skill that is needed to succeed on the job market, and while this aspect is not mentioned in relation to the future in any other focus group, the sentiment is shared by others, particularly by those participants who view English as having become 'normal' as indicated by this statement by Franz M.:

> **Franz M.**: [...] und es ist auch wichtig es [Englisch] zu können, weil es irgendwie normal geworden ist, weil es fast jeder kann (SC01: 12)
>
> **Franz M.**: [...] and it is also important to know it [English] because it has somehow become normal, because almost everyone knows it

Finally, two further reasons that are given for the significance of English in Austrian adolescents' lives are the use of Anglicisms in German and the influence of English on youth language. Two participants comment that many words in German are actually English loanwords and that knowing English helps to understand these more easily (Lia SA02: 17, DJ SF01: 62-64). Walküre provides the example of her grandmother passing billboards and wondering what they say (SF01: 70-73) implying that because of the large number of Anglicisms in general and in

advertisements in particular older people, who often do not speak English well, find themselves at a disadvantage. While this example may be slightly exaggerated, she then goes on to explain that English has become an integral part of Austrian youth language:

> **Walküre**: Und auch halt weil, weil, weil Englisch ja mittlerweile schon relativ viel in die Jugendsprache also reingeschwappt ist,
> **SX-m**: Ja.
> **Walküre**: dass halt viele auch irgendwelche englischen Begriffe benutzen und dann steht man so und unterhält sich mit jemanden und hat eigentlich keine Ahnung was der grade gesagt hat, das ja (.) ist dann halt auch irgendwie anstrengend, wenn man nachfragen muss. (SF01: 75-78)
>
> **Walküre**: And also well because, because, because by now quite a lot of English has flown over into youth language,
> **SX-m**: Yes
> **Walküre**: so that many just use some English terms and then you are standing there talking to someone and you actually have no idea what that person just said, that yeah (.) that's also annoying somehow, if you have to ask.

This perception is echoed by Johannes, who also comments that English terms and abbreviations have been included in youth slang and that it would seem strange if a young person did not understand these.

To sum up, the participants attribute a high degree of significance to knowing English and provide examples as well as reasons for its role in their lifeworlds. They emphasize that English is of importance for young Austrians because it is needed for leisure time activities and for staying up to date, and because of the fact that many expressions have become part of Austrian youth language. Considering a wider context, participants argue that knowing English is essential because it is the 'universal language' used in international conversation and thus useful for holidays or stays abroad as well as their future careers. In short, knowledge of English is perceived as a highly valuable resource for young Austrians and even more so than other languages, as discussed in the next section.

7.1.1 Comparisons of English with other languages

Related to the topic of the significance of English for young Austrians is the comparison of English with other languages. Within this data-driven category (see section 5.4.6) three main aspects surface in the focus group interviews: evaluations of the importance of English with regard to other languages, evaluations of the aesthetic qualities of English in comparison to German and other languages, and evaluations of the influence English exerts on other languages, again especially on German.

The first aspect is connected most closely to the topic of significance discussed in the previous section, as several participants use comparisons with other languages to illustrate the significance of English in their lives and consistently describe English as the most important language after their first language(s). In statements concerning the other two aspects, aesthetic qualities and influence on other languages, English is commonly contrasted with German. In almost all focus groups participants comment on the aesthetic qualities of German and English at some point, describing German as long and complicated (John M. SC01: 201), hard and staccato

(Paul SC01: 207-211), or simply "lame" in connection with specific EE activities (Kirito SE01: 245). In contrast, English is perceived as a more beautiful language (Anna SD01: 36), sounding better (Lia SA02: 155), and allowing more ways to express oneself (Kira, Lia and John W. SA02: 35-41). Such positive evaluations of English also play a role in participants' decisions to use it for their spare time activities (see section 7.2).

Lastly, the influence of English on other languages, in this context particularly on German, emerges as a controversial issue in the interviews. Some students describe the phenomenon of lexical borrowing from English in neutral terms; for instance, those who argue that knowing English is important for young Austrians because many terms have already entered the German language (see Walküre's quote in section 7.1) portray lexical borrowing as a fact and do not provide any evaluation as to whether they view this development as positive or negative. In contrast, other participants regard Anglicisms and the use of English words and phrases in German as clearly negative and become quite emotional about it. The following extract from focus group SA02 about using English words while speaking German illustrates these two opposing views represented by Keanu and Niall:

> **Keanu**: Ja, solche Wörter ‚Das ist so <ENGLISH> nice </ENGLISH>', ach, da krieg ich manchmal solche Wutanfälle.
> **SS**: @
> **Keanu**: Weil das, ich geh auf der (.) irgendwelche Freunde von mir sagen plötzlich ‚Ja urnice', also keine Ahnung, tschuldigung, aber wir sind in Österreich, ich hab nichts gegen englische Sprache @, aber entweder man spricht ganz auf Englisch oder ganz auf Deutsch. Aber dieses Anbauen, ich find das das das macht ur die deutsche Sprache weg, irgendwie. Keine Ahnung.
> **Niall**: Was das macht, ist jetzt eigentlich <ENGLISH> off-topic </ENGLISH> aber es ist ja eigentlich
> **Interviewer**: Wurscht, wir dürfen auch <ENGLISH> off-topic </ENGLISH> gehen.
> **Niall**: es ist in jeder Sprache so, dass es durch andere Sprachen stark beeinflusst wird, das Worte übernommen werden, dass ja, diese Anglizismen sind jetzt nicht wirklich was Neues.
> **Keanu**: Ja, aber was mich oft auf-,
> **Niall**: Das sag ich jetzt nicht nur, weil ich oft <ENGLISH> nice </ENGLISH> sage und mich grade schuldig fühle (SA02: 127-134)

> **Keanu**: Yes, such words 'that's so <ENGLISH> nice<ENGLISH>', ah, sometimes that makes me really angry
> **SS**: @
> **Keanu**: Because it, I walk on the (.) some friends of mine suddenly say 'Yeah, uber-nice', well no idea, excuse me, but we are in Austria, I have nothing against the English language @, but either you speak completely in English or completely in German. But this attaching [of English and German words], I think that that that totally takes the German language away, somehow. No idea.
> **Niall**: What does, it's actually <ENGLISH> off-topic </ENGLISH>, but actually it is
> **Interviewer**: Fine, we can go <ENGLISH> off-topic </ENGLISH>.
> **Niall**: it's like that in every language that it's strongly influenced by other languages, that words are taken over, that, well, these Anglicism really aren't anything new.
> **Keanu**: Yeah, but what really upsets me-
> **Niall**: I am not just saying that now because I often say <ENGLISH> nice </ENGLISH> and I feel guilty right now.

Interestingly, the word *nice* came up in two other focus groups as well: Kirito (SE01: 125-130) uses it as an example to explain to another participant what is meant by 'using English words in other languages' and Paul (SC01: 1) says he often uses it, but he sees it as an already Germanized

word form. Words like *nice* could be characterized using Elisa's lexical creation *"Parasitenwort"* [a parasite word] (SG02: 324), although she uses this expression to refer to certain English (swear)words that she constantly sees online and frequently uses herself and to her using quotative *like* even in German, which she clearly evaluates negatively.

While there clearly is no consensus among participants as to whether code mixing and lexical borrowing are a long-standing fact or an undesired phenomenon, the majority agrees that English plays an important role in their daily lives and more so than other foreign languages, although the degree of significance attributed varies. The reasons given for the importance of English for young Austrians range from its use as an international lingua franca, which has perceived implications for both stays abroad and participants' future careers, over its use for leisure time activities and the need to stay up to date, to its influence on Austrian German and youth language. Overall, it became clear that knowing English is regarded as a valuable and sometimes necessary resource by participants. As one of them expressed it:

> **Mito**: Ich hab bis jetzt noch nie irgendeine andere Sprache gebraucht als Englisch so in meinem Alltagsleben. Halt @ wir lernen Englisch, Französisch und Latein und bis jetzt hab ich immer nur Englisch gebraucht. (SD01: 12)

> **Mito**: Until now I have never needed any language other than English in my everyday life. I mean @ we learn English, French and Latin but up to now I have only ever needed English.

7.2   Descriptions of extramural practices

Figure 7.2: Word cloud of the 100 most frequent types in the main category DESCRIPTION OF EE

Having found that students regard English as important for their everyday lives, the next step is to see how it features in their leisure time activities. This section sets out to answer the following two research questions *How do participants describe their own EE practices?* (RQ 5b) and *How do participants interpret and evaluate the results of the quantitative part of the study?* (RQ 5c).[199] We will begin with the latter before exploring participants' accounts of their current activities in section 7.2.1. Subsequently, comments on the amount of time spent with EE (section 7.2.2) and

---

[199] In this section information will be provided on how the interviewees evaluated the quantitative results concerning the most popular EE activities and amount of time spent with EE, whereas the evaluation of results relating to vocabulary learning strategies will be discussed in section 7.4.1.

reasons for using English in their spare time (section 7.2.3) are presented. Additionally, the types of English participants encounter during their free time are discussed in section 7.2.4.

7.2.1 Current EE activities

In the interviews participants were presented with preliminary quantitative results in form of visual input. Pie charts were used to present the eight most popular EE activities, which more than 50% of the students participating in the quantitative strand reported doing at least a few times a week (see section 5.4.2 and Appendix B).[200] Participants were asked whether they found any of these findings surprising or whether they had any other comments. Since the quantitative results have been described in detail in Chapter 6, in the description of the qualitative outcomes below, emphasis will be given to those results that complement and enhance the quantitative findings.

Overall, students in all focus groups agree that the quantitative results are not unexpected. In addition, the EE activities participants named when asked to describe the significance of English in their everyday lives (see section 7.1) largely coincide with the most popular activities identified in the quantitative strand. Participants explain that the popularity of these EE activities is rather obvious (Lukas SG02: 67) or quite predictable (Maria SC01: 72), but find different details surprising.

The fact that music is the most frequent EE activity is regarded as absolutely obvious since most music they listen to is in English (e.g. SD01: 80-86) and the finding that series and films are frequently watched in English by Austrian teenagers is evident to participants as well. Discussions of films and series in the focus groups frequently centred around the question which of the two is being watched more often and suggest that the time factor, i.e. the length of an episode or film, is not decisive. Arguments brought forward include that series are seemingly shorter and thus easier to watch on a daily basis, but students report frequently watching several episodes at once, which then actually takes longer than a movie (John M., Maria & Paul SC01: 58-69, Johannes & Lukas SG02: 98-114). The following extract exemplifies an example of such binge watching:

> **Lukas**: Zum Beispiel die Serie <ENGLISH> Game of Thrones </ENGLISH> hab ich irgendwie später angefangen, also da waren schon die ersten vier Staffeln draußen.
> **Interviewer**: […]
> **Lukas**: Aber, es ist es ist halt so, da dauert ja eine Folge meistens eine Stunde,
> **Johannes**: Ja.
> **Lukas**: eine Stunde und zwanzig irgendwie sowas halt und ich weiß schon, das hab ich irgendwie in einer Woche da durchgesuchtelt einfach weil ich wissen wollte wie's weitergeht.
> **Interviewer**: Wow. Hallelujah.
> **Johannes**: Eine Woche ist schon ziemlich (schnell).
> **Lukas**: Ja. (SG02: 116-123)

---

[200] These activities are listening to music, watching video clips, watching films, watching series, reading in social media, using English words in other languages, using search engines, and using apps (see section 6.3.1).

> **Lukas**: For example, I somehow began later with the series <ENGLISH> Game of Thrones </ENGLISH>, I mean by then the first four seasons were out
> **Interviewer**: [...]
> **Lukas**: But it's, it's just, one of the episodes usually lasts about an hour,
> Johannes: Yes.
> **Lukas**: An hour and twenty or something like that and I know that I somehow binged that in one week just because I wanted to know what was going to happen next.
> **Interviewer**: Wow. Hallelujah.
> **Johannes**: A week really is pretty (fast).
> **Lukas**: Yeah.

Paul (SC01: 87-97) describes experiences with series similar to Lukas' and reports watching a lot of one series over a short period of time, which is then often followed by a month or more of not watching anything because he has not found another show that interests him. In this regard, the qualitative data add a new perspective in that they show that patterns of consumption for 'online TV' are not stable over time. It appears that binge watching followed by rather extensive periods of viewing no or less series is relatively common and that the quantitative results in the questionnaire and language diary thus have to be regarded as a snapshot of what students were doing at the time.

In addition to films and series, participants also discuss watching online videos, which they usually view on YouTube or come across on social media platforms like Instagram (Anna, Emma & Susi SD01: 92-99, DJ, KingKong & Walküre SF01: 94-100). Here, more information could be collected on popular genres: these include clips by English-speaking YouTubers, comedy clips and so-called let's plays, where viewers watch a gamer play a game with running commentary, (Johannes & Lukas SG02: 129-144, Jane SC01: 435) as well as videos about celebrities (Keanu SA02: 54).

In contrast to the quantitative data, gaming plays more of a role in the focus group data: four male students, John W. (SA02: 221), Paul (SC01: 461), Kirito (SE01:36-48) and Lukas (SG02: 13-17), appear to play rather a lot and they explicitly mention using English to communicate with others during gameplay. Lukas emphasizes that collaboration with other gamers is key if one does not want to solve everything on one's own and therefore one has to communicate with others. For this reason, Lukas and Paul speak to fellow gamers using headsets and VOIP services, while John W. and Kirito prefer chatting. However, all boys agree that using English is a must for in-game communication, as exemplified by the following statement:

> **John W.:** Ja, ahm, ich schreib, ich schreib eigentlich auch nicht so wenig, weil ich hab ahm so, weil ich spiele relativ gerne Computer und relativ viel, ahm [clears his throat] und ich schreib auch da mit mit Kontakten, die halt nicht immer auch nicht immer Deutsch können und dann muss ich eben gezwungenermaßen auch Englisch schreiben. (SA02: 221)
>
> **John W.:** Yes, ahm, I write, I actually do not write so little either because I have, ahm, so because I quite like playing computer games and quite a lot, ahm [clears his throat] and then I write with with contacts, who don't always know German and then I am also forced to write in English.

In contrast, participants explain that in single-player games one can usually choose the language and argue that in this respect single-player games are comparable to series or films, as there is a choice between using the original English or a dubbed German version (SC01: 475-485). In line with the quantitative data, these results show that while there are few gamers in the sample,

those who play games do so a lot. The focus group data further show that its lingua franca status in multiplayer online games is a major contributing factor to the use of English in gaming.

The productive use of English during leisure activities is also discussed in the focus groups, but interestingly almost exclusively in relation to writing, whereas speaking is hardly mentioned at all. This omission can be seen as another indication that English speaking is a marginal activity for the majority of Viennese teenagers. With regard to writing in general, there is a consensus across several groups that teenagers hardly ever write in English on social media or other online platforms (SC01: 147-157, SD01: 128-131, SF01: 115-121, SG02: 89-96), aptly summarized by Niall:

> **Niall**: Also ich versuche nicht aus dem Weg zu gehen, aber es passiert halt kaum. (SA02: 229)
>
> **Niall**: Well, I don't try to avoid it, but it just hardly ever happens.

Some students like Paul (SC01: 147) think that writing in English takes too much effort for many young Austrians, while Susi (SD01: 131) argues that many do not see a need to comment a lot online in general. Both Johannes (SG02: 90) and Walküre (SF01: 130-135) explain that it also depends on the context: if the language of the video or article is German, one would also use German in comments; but if the original post or previous comments mostly use English, they would write in the same language. The only participants who report frequently using English online again use it because of its lingua franca role: Jane's (SC01: 149-153) group of friends includes both Russians and Austrians, which is why she uses English for almost all her posts on social media.

Interestingly, several other participants point out that while they do not frequently post or comment in English publicly, they use it for private (online) communication with friends. DJ brings forward an argument similar to Susi's above, stating that he usually does not write comments, but he uses English to chat with both German- and non-German-speaking friends:

> **DJ:** Also ich bin zum Beispiel so einer in den sozialen Medien (.) ich ich bin nicht so einer, der jetzt jeden Tag auf sozialen Medien ist, außer YouTube, aber da schreib ich auch keine Kommentare,
> **Interviewer**: Mhm.
> **DJ**: fast also eigentlich gar keine,
> **Interviewer**: Ja.
> **DJ**: halt ich schreib mit meinen Freunden, mit meinen Freunden schreib ich sehr viel auf Englisch.
> **Interviewer**: Mhm.
> **DJ**: Aber (.) also ich hab auch englischsprachige Freunde mit den- mit denen schreib ich auf Englisch und außerdem mit Freunden zum Beispiel, keine Ahnung wieso, wir rutschen auf ins Deutsche ins Englische rüber. (SF01: 115-121)
>
> **DJ**: So for instance I'm not one of those on social media (.) I I am not one of those who is on social media every day, except YouTube, but there I also don't write any comments.
> **Interviewer**: Mhm.
> **DJ**: almost or actually none at all
> **Interviewer**: Yes.
> **DJ**: I write with my friends, with my friends I write a lot in English
> **Interviewer**: Mhm.
> **DJ**: But (.) well I have English-speaking friends with tho- with those I write in English and in addition with friends, for example, no idea why, we just slip from German into English.

Keanu (SA02: 211), Anna (SD01: 25-27) and Jane (SC01: 463-465) also have international friends who do not speak German and hence use English to message them and Walküre (SF01 30-34) uses it in her community of online friends. Maria (SC01: 471-474), Pinguin and Kirito (SE01: 296-306) use English with relatives who do not live in Austria because they do not know their parents' native languages well enough for communication with their extended families. Like DJ in the quote above, Pinguin (SE01: 165-171) also says that he predominantly uses English with an Austrian friend because they just happen to chat in English all the time. The majority of participants however concur that rather than for writing they use English for reading in online environments and their comments suggest that they encounter written English in a variety of formats online.

Discussions of online reading habits indicate different reading behaviours: some mainly encounter written English in memes or jokes on 9Gag or Instagram (John SE01: 140-146, KingKong SF01: 18-22), whereas others report reading longer texts such as news articles or information texts on Wikipedia (Mito SD01: 71-78; Pinguin, Karl and Kirito SE01: 140-146). However, students do not only read for information, but also for pleasure: Marie (SG02: 489) is a big fan of the app Wattpad, where one can access other users' stories for free as well as react to them, and Kirito (SE01: 4-12) reports reading English stories on the internet for a considerable amount of time every day. Overall, it appears that a lot of reading in English happens in online contexts, but it is not limited to these. Some participants report reading books in English: Jane (SC01: 389-393) likes graded readers because she does not enjoy reading as much if the language level is too difficult for her, while John M. (SC01: 399-405) takes great pleasure in reading original versions despite sometimes encountering difficulties with the language. Finally, Susi (SD01: 218-220) expresses a general preference for reading books in English rather than in German because it also presents an opportunity for developing her language competence.

In addition to providing further insights during their discussions of EE activities among Viennese teenagers, students also arrive at two conclusions that mirror my own analysis in section 6.3.1. Looking at the visualization of the most frequent EE activities, Johannes (SG02: 69-71) notes that almost all of the popular activities are done online, while much later in the same focus group Louise realizes that most popular EE activities only involve receptive language skills:

> **Louise**: Das ist ja fast nur passiv. Nur die beiden sind aktiv.
> **Interviewer**: Ja. Genau.
> **Louise**: Ja. Man man macht mehr, also man hat einfach mehr passiv Englisch. (SG02: 804-806)

> **Louise**: That's almost only passive. Only those two are active.
> **Interviewer**: Yes. Exactly.
> **Louise**: Yes. You do more, well, you just have more passive English.

In another group, John M. (SC01: 145) expresses a similar thought when he comments that the most frequent activities involve little speaking or using English. These instances do not only show that qualitative follow-up interviews can be usefully integrated with quantitative results, but also showcase teenage participants' ability to interpret data and draw their own conclusions,

which supports the usefulness of techniques such as member checking (Creswell & Miller 2000). However, similar to the quantitative results, the qualitative data also show that in addition to the more frequent activities there is a broad range of interests among individual adolescents as four participants describe highly specialized niche activities, which are discussed in more detail below.

Kira is the only girl who describes such a specialized activity; she generally appears to engage with English a lot: she reads novels, likes to watch British series such as *Sherlock* and *Doctor Who* and frequently communicates with Austrian friends who moved to the US some years ago. In addition, she enjoys immersing herself in online fan worlds, in the following segment she describes an app she has recently discovered:

> **Kira**: [...] aber jetzt hab ich eine App, da kann man halt, es ist wie ein Art Facebook nur halt auf ein bestimmten Film oder Serie oder Buch fixiert.
> **Interviewer**: Mhm.
> **Kira**: Zum Beispiel gibt's das für Harry Potter und da les ich jetzt halt täglich (.) sehr viel und das sind aber alles Mögliche, das sind ahm Quizzes, das sind irgendwelche Fragebögenart, dann gibt's eben auch wieder solche Blogs und so etwas [...] (SA02: 99-101)

> **Kira**: but now I have an app, where you can, it's like a sort of Facebook, but focusing on a specific film or series or book.
> **Interviewer**: Mhm.
> **Kira**: For example there is one for Harry Potter and there I read every day now (.) a lot and there's all sorts of things, there are ahm quizzes, there are some questionnaires, then there are blogs and things like that.

However, Kira does not only consume fan culture, she also actively contributes stories, poems or comments, most of which are in English. Creative writing certainly is a specialized EE activity in which most Viennese teenagers do not engage; moreover, it is an activity that requires an advanced level of language proficiency. Kira thus is a good example of how a special interest may lead to intensive out-of-school use of English, which is likely to aid the development of English language skills.

Another niche activity is computer programming, which Mito frequently does in his free time; in particular, he programs extensions for the game *Minecraft*. While programming itself is not necessarily connected to the use of English, researching the problems he encounters or reading specialized texts about programming definitely is, as Mito explains in the following extract:

> **Mito**: Ja, also in meiner Freizeit tu ich halt oft so Programmieren und so,
> **Interviewer**: Uhh.
> **Mito**: und wenn man da halt irgendwas nachgoogelt, dann kann man das fast nur auf Englisch machen,
> **Interviewer**: Ja, das glaub ich.
> **Mito**: und alles was man halt hier irgendwo auch liest oder so, liest man halt auch immer auf Englisch,
> **Interviewer**: Das ist ziemlich cool.
> **Mito**: weil es im Deutschen nicht viel gibt. (SD01 14-20)

> **Mito**: Yes, well in my free time I often do programming and such.
> **Interviewer**: Uhh.
> **Mito**: and when you look something up on Google for that, then you can pretty much only do it in English.
> **Interviewer**: Yes, I believe that.

> **Mito**: and everything that you read somewhere or so, you also always read it in English
> **Interviewer**: That's pretty cool.
> **Mito**: because there's not much in German.

Since much more specialized information on computer programming is available in English and Mito has to use it to develop the expertise needed for his hobby, he encounters English in a highly technical context. In a similar fashion, Pinguin also needs English to find information for his hobby as he likes to produce video clips in his free time and the information he needs to solve problems is usually only available in English. However, Pinguin also points out that the medium of online video itself is connected to English. While we do not know what type of videos Pinguin likes to make and whether they involve voice-overs or language in any other way, it is safe to assume that he frequently encounters English when engaging in his hobby and is also likely to use it productively to some extent.

Unlike the examples of niche activities so far, the last one does not take place in an online context, but is a much more physical activity:

> **DJ**: Ahm, für mich ist Englisch eigentlich sehr sehr wichtig, weil ich spiel in einem Verein, Basketball,
> **Interviewer**: Ok.
> **DJ**: und mein \<ENGLISH\> Coach \</ENGLISH\>, also mein Trainer, er spricht die meiste Zeit Englisch und (sagen) halt Englisch auch so, wir verständigen, wir verständigen uns auch alle mit ihm über über über die englische Sprache,
> **Interviewer**: Mhm.
> **DJ**: und zum Beispiel auch Spielzüge, zum Beispiel Taktiken, sagen wir auch auf Englisch auf.
> **Interviewer**: Mhm.
> **DJ**: Also,
> **Interviewer**: Das heißt du sprichst eigentlich nur Englisch im Basketballtraining?
> **DJ**: Genau. (SF01: 2-10)

> **DJ**: Ahm, for me English is actually very very important because I play basketball in a club.
> **Interviewer**: Ok.
> **DJ**: and my \<ENGLISH\> coach \</ENGLISH\>, so my trainer, he speaks English most of the time and (we say) English too, we communicate, all of us we communicate with him in the English language,
> **Interviewer**: Mhm.
> **DJ**: and for example the plays as well, for example the tactics, we also say them in English.
> **Interviewer**: Mhm.
> **DJ**: So,
> **Interviewer**: That means you actually only speak English during basketball training?
> **DJ**: Exactly.

DJ's main extramural English input is his basketball training and the communication with his coach. In other sections of the focus group interview (SF01: 161, 175, 262), DJ explains that he joined the basketball club three years ago and really trains extensively with his team: he goes to practice, which sometimes lasts up to four hours, five times a week. He explains that speaking English to his trainer, who is US American, and his teammates has led to a vast improvement in his level of English, a point which will be taken up again in section 7.3. Unlike the more typical EE environments of Viennese teenagers, DJ's is characterized by productive language use. Because of basketball, which takes up a lot of his free time, DJ reports not watching series, but he likes listening to rap music and going to the English cinema and often finds time to watch

short YouTube clips. Compared to other participants he therefore does not spend as much time on the most popular EE activities, but he regularly uses it productively in oral communication.

These examples of four specialized niche activities from different focus groups show that for different individuals the variety of leisure time pursuits involving English is much greater than indicated by the most common EE activities identified in the quantitative results. However, in the same way, there is also great variation in relation to the amount of English students use in their spare time: some, like the four cases presented above, have extensive contact with English, whereas others use it very little. One limitation of the focus groups thus is that most of the participants use a lot of EE with 19 out of 30 participants reporting that they use English more in their free time than during school lessons (see section 5.4.4).

7.2.2 Time spent with extramural English

In the focus group interview students were asked to comment on another finding of the quantitative strand: the amount of time spent with EE, which according to the EEOLD data (see section 6.3.2) was approximately four hours a day on average. However, even before this point in the interview, many participants commented on the frequency of encountering English; in each focus group one or more participants explicitly mentioned using English outside school on a daily basis for a variety of purposes.

When confronted with the mean EE time of approximately four hours, participants' spontaneous reactions are quite different: at first many students express disbelief because four hours are a large amount of time (e.g. Maria & Jane SC01: 232-233, Franz SE01: 189-191), but upon consideration and discussion among each other, most of them come to the conclusion that this estimate could actually be accurate (e.g. Maria & Jane SC01: 253-260, Susi SD01: 164-167, Franz SE01: 200, Elisa & Johannes SG02: 180-181).[201] In addition, seven participants immediately affirm that this estimate seems plausible to them (Kira SA02: 236-240, John W SA02: 236-275, Pinguin, Karl & Kirito SD01: 185:187, Walküre SF01: 140-145) and three claim that for them it is probably even too low (DJ & KingKong SF01: 139-145, Lukas SG02: 151-159). While an ever higher amount of time is surprising, it is not entirely implausible for these three students: as described in section 7.2.1, DJ spends many hours a week training with his English-speaking basketball team, KingKong appears to spend most of his time online although it does not become quite clear what he actually does on the internet, and Lukas (SG02: 151-159) is a gamer who frequently plays multiplayer online games. In contrast, three students explain that it is definitely not true for them (Keanu & Lia SA02: 242-256) or argue that they do not believe it is true for the average Viennese teenager (Anna SD01: 137-139).

In sum, participants' evaluations of the amount of time spent with English outside school range from total disbelief to even higher estimates, but the majority of participants decided it was plausible once they had had time to reflect on it and discuss it.

---

[201] Out of 25 participants who gave an evaluation of the average EE mean time based on the online language diary, 12 students explicitly expressed a conclusion that on closer thought this estimate was plausible.

The following extract represents a typical discussion about the time spent with EE:

> **Paul**: Vier Stunden klingt jetzt eigentlich wirklich wirklich viel
> **Maria**: aber es könnt hinkommen.
> **Paul**: Genau, <13> es könnt </13>
> **John M.**: <13> xx einfach aufgeteilt, </13>
> **Jane**: Mit Musik und so,
> **Maria**: Man schaut sich mal so nen Film an, das sind schon mal zirka zwei Stunden,
> **Franz M.**: Ja.
> **Maria**: <14> und dann (hört man noch) Musik und das passt. </14>
> **Paul**: <14> Vielleicht nicht Sachen wo man bewusst Englisch benutzt, vielleicht nicht direkt, also wo man nicht aktiv Englisch redet, aber was halt </14> [...]
> **Paul**: Also ich glaub Musik nimmt da wirklich einen großen Teil ein,
> **SS**: Ja.
> **Paul**: weil weil's einfach etwas ist, weil wenn du dir einfach denkst ,ja, vier Stunden Englisch, so viel Zeit hab ich ja gar nicht zum (jetzt vier Stunden mich) hinsetzen Englisch machen,
> **Interviewer**: Ja.
> **Paul**: aber Musik ist ja das was zwischendurch ist und,
> **John M.**: Ja.
> **Maria**: man kann ja was anderes währenddessen machen (SC01: 255-274)

> **Paul**: Four hours actually sounds really really a lot
> **Maria**: but it could be right.
> **Paul:** Exactly <13> it could </13>
> **John M.**: <13> xx just split up </13>
> **Janes**: With music and so on,
> **Maria**: You just watch a movie, that's already about two hours
> **Franz M.**: Yes
> **Maria**: <14> and then you also listen to music and that's it </14>
> **Paul**: <14>Maybe not things where you use English consciously, perhaps not directly, so where you do not speak English actively, but that just </14> [...]
> **Paul**: Well I think music really takes up a large part of it.
> **SS**: Yeah.
> **Paul**: simply because it's something, because you simply think 'Ok, four hours English, I don't even have that much time to sit down and do English'
> **Interviewer**: Yes.
> **Paul**: but music is something that happens in between and,
> **John** M.: Yes
> **Maria**: you can do something else in the meantime

Similar segments in which students try to estimate the time they usually spend on different activities are found across all focus groups and music emerges as an important time factor in these. Franz (SE01: 197-200) explains that many teenagers constantly listen to music and Elisa (SG02: 173-175) explains that she purposefully takes a longer way when she has to walk somewhere so she can listen to more songs. This is also an example of the importance of 'doing EE activities on the move': through smartphones many types of (English-language) media are available to teenagers wherever they are. This kind of mobile engagement with EE should not be disregarded as several participants mention listening to music (Emma SD01: 152, Elisa SG02: 175), watching video clips (DJ SF01: 173), using social media and reading online (Paul SC01: 280-282), or reading books (Kira SA02: 238) while going somewhere on foot or public transport.

In addition, two further aspects are mentioned in the quote above: first, Paul clarifies that teenagers are unlikely to actively or even consciously use English for four hours and secondly, at the end of the quote Maria mentions the possibility of doing several things simultaneously.

The possibility of two or more EE activities done simultaneously could partly explain the high estimate of EE time found in the EEOLD and therefore this idea was also discussed during the focus group interviews, if time allowed. Some participants in group SD01 (168-175) think that it could be the case that students who filled in the online language diary counted time spent on parallel activities twice, whereas in group SA02 (261-272) participants argue that that was unlikely because they themselves did not do so and there are few EE activities which can really be done simultaneously.

In sum, the majority of participants find the mean EE time of approximately four hours suggested by the EEOLD data plausible. They emphasize the role of music, which in their opinion could account for a rather large part of this exposure to EE, as well as the role of mobile media usage. Additionally, some participants point out that teenagers are unlikely to consciously use English for four hours a day and there is no agreement as to whether doing EE activities simultaneously could account for the high estimate of time spent with English per day.

7.2.3 Reasons for using English

In addition to describing their current and past EE activities, participants also provide reasons for using English in their free time. To a certain extent these overlap with the reasons given for the significance of English in young Austrians' everyday life discussed in section 7.1, especially with the idea that English is needed for leisure time activities and for international communication. The reasons for using English for leisure time pursuits found in the data are summarized in Table 7.2 and described in more detail below.

Subcategory	Number of coded segments	Number of participants
Original version	13	13
English is cool or beautiful	12	10
Many things just are in English	11	8
Friends or family abroad (who don't speak German)	10	10
Wider pool of information	9	10
Easier to express oneself in English (more choice)	6	7
International communication (while gaming)	5	5
Availability and being up-to-date	4	5
Pop culture is influenced by English-speaking countries	3	3
Not having to switch language	2	3
Better content	2	2
Used as a secret language	2	2
Intentional learning	2	1

Table 7.2: Frequencies of occurrence for the subcategories of REASONS FOR USING ENGLISH

One of the main reasons given for using English for spare time activities is that the English language version of films, series or books often is the original. Many participants express a dislike for the dubbed versions of films and series because the voices in the German versions are simply 'wrong' (Vanessa SD01: 22, Walküre & DJ SF01: 26-29) or sound artificial to them (Kira

SA02:18-22) and they argue that once you have heard the original you cannot go back (John M. & Maria SC01: 171-175). Moreover, John W. (SA02: 25-27, 165) points out that jokes are always better in the original, especially if they involve wordplay, and that the original version is what a film or series was meant to be like. Similar to films or series, the original version is also better for single-player games because dubbed versions often lack quality according to Paul (SC01: 483). Yet, English originals are not just more attractive for audiovisual media, but also for books: Paul (SC01: 395-397) expresses an interest in re-reading books he liked in German in English to see what the original version is like and Susi (SD01: 50) prefers reading books in English if they were first published in this language.

Other reasons discussed by participants echo their arguments for the overall significance of English in their lives (see section 7.1): the aesthetic qualities ascribed to English and its role as 'universal language' which influences the availability and quality of content as well as communication with international family and friends (see description in section 7.2.1). Several students explicitly refer to aesthetic qualities (Anna SD01: 27-36; Walküre SF01: 30, 232) or mention a preference for English because they like it (e.g. Lia SA02: 155-160, Paul & John M. SC01: 6, 33-35, Pinguin SE01:179). Besides, a few students think that sometimes it is easier to express oneself in English: in group SC01 (944-956) the students discuss that it is much easier and more striking to swear in English, while in group SA02 (40-41) two girls argue that English offers a greater variety of expression in comparison to German, in particularly when it comes to wordplay.

With regard to availability, participants argue that many things – be it music, films, series, videos, texts or games – are just available in English, especially on the internet (e.g. Anna SD01: 117). Several students mention that the content they encounter on online platforms like 9Gag (KingKong SF01: 18-22), Instagram (Susi SD01: 99, Marie SG02: 790-792), Tumblr (Pinguin SE01: 176) or YouTube (Paul SC01: 167-169) is mostly in English. Hence, in addition to the ubiquity of English-language music, the exclusive availability of online content in English seems to be a strong reason for teenagers to engage with the language during their free time. Another reason for engagement with English-language media could be the orientation towards US American pop culture in Western and Central European countries: three girls in different focus groups (Keanu SA02: 51-58, Walküre SF01: 88-90, Elisa SG02: 203-209) suggest that Austrian teen culture is strongly influenced by the US and express more interest in American music and stars than in their Austrian and German counterparts. Moreover, media content may not be exclusively available in English, but it is usually published a long time before the dubbed versions (see also section 7.1). Students in three focus groups agree that not wanting to lag behind with a TV series, for instance, is a reason for watching it in English (Paul SC01: 48-49, Vanessa & Anna SD01: 51-59, Karl SE01: 31).

In addition, the quality of content available in English also plays an important role: Emma (SD01: 96-98) states that in her opinion English-language content is often better, particularly when it comes to user-generated content like YouTube videos, and Susi (SD01: 24) argues that there is a

much wider selection of online media available in English. However, the fact that more and more diverse content is accessible through English is not only important for entertainment, but also for finding information. Participants in five out of six focus groups state that they use English to search for information online (Kira & Niall SA02: 112-115, Paul, John M. & Maria SC01: 281-226, SD01: Mito 71-78, Pinguin & Franz SE01: 148-156, KingKong & Walküre SF01: 103-107) and use sites such as Wikipedia (John M. & Maria SC01: 224-225) in English rather than in German. Many refer to the wider pool of information available in English as a reason, similar to Niall in the quote below:

> **Niall**: Gut. Ähm, ich find Englisch für mich mehr wichtig (.) ist es wichtig, weil es einfach (.) du viel mehr Möglichkeiten hast auch Themen zu finden. Also, we- we- wenn du i- ir- irgendetwas über ein Thema herausfinden willst, dann hast du nicht nur den deutschen Sprachpool, sondern den englischen, der irgendwie wahrscheinlich eh der größte ist von allen (1) vielleicht abgesehen vom chinesischen, aber (2) sonst ja [...] (SA02: 22)

> **Niall**: Okay. Ehm, I think English is more important for me (.) it's important because there is just (.) you have so many more possibilities to find topics. So, if if if you want to find out anything about a topic, then you don't just have the German language pool, but the English one, which somehow probably is the largest of all (1) perhaps apart from the Chinese, but (2) well yeah [...]

One last aspect related to using English-language media is mentioned by two participants in group SE01: statements by Karl (SE01: 115-121) and Kirito (SE01: 181-183) indicate that there is a type of chain effect when it comes to their EE activities. The boys explain that once they do an activity in English and they need to do something connected to it, like for instance find further information, write a comment or respond to someone, they automatically do that in English as well because they do not want to translate into another language or "think twice" as Karl puts it. Although only these two participants explicitly report such behaviour, their statements could indicate more widely adopted practices.

Finally, two further reasons were given for using English that are not related to typical EE practices. These are intentional learning (George SE01: 52-57) and the use of English as a secret language to keep parents or younger siblings from understanding conversations in the family (Kirito SE01: 311-315, Elisa SG02: 213-215).

On the whole, the main reasons participants offer for their engagement with English during their leisure time are a preference for original versions, the exclusive availability of online content in English or the delay in the availability of dubbed versions, and the aesthetic qualities ascribed to the English language. In addition, they argue that English is needed for international communication and keeping in touch with non-German-speaking friends and relatives and gives access to a wider pool of information, especially online. Other reasons given relate to the better selection of content coupled with a general orientation of Austrian teenagers towards US American pop culture, the possibilities of expression in English and a chain effect, where doing one activity in English leads to a related EE activity.

7.2.4 Types of English encountered by participants

A topic related to participants' EE activities is the types of English they come across both inside and outside the school context. This topic sometimes emerged during interviews, but if time allowed, students were also explicitly asked about the kinds of English they encounter through their EE activities. In response they described their contact with and opinions on different varieties of English as well as different registers.

Several participants report experiencing different varieties of English through EE input (SG02: 283-290) and some mention that they like listening to different accents and try to differentiate between them, referring to Australian and Scottish English, for example (Jane & Paul SC01 182-196). With regard to the variety of English they encounter most frequently, most participants named American English (Anna SD01: 370, Franz & Pinguin SE01: 212-225, DJ SF01: 212). As reasons they suggest that many series are produced in the US (Franz SE01: 223:225) and that there are more speakers of US American English worldwide (Mito SD01: 388-390). In contrast, a few students express a preference for British English: Karl (SE01: 216- 218) thinks he encounters British English more often and Walküre (SF01: 214-218) likes British English better and thus intentionally watches more British YouTubers, but she points out that with series the choice is more limited because the variety used depends on whether they were produced by the BBC or by American TV stations (SF01: 224-226).

In addition to varieties, participants also discuss other types of English they encounter outside school, which are mainly characterized by differences in register. While DJ and KingKong (SF01: 184-210), who like listening to hip hop, report often encountering "*Straßenenglisch*" ["street English"] and slang terms, other examples show that students do not just engage with everyday high frequency vocabulary in their EE activities: John M. (SC01: 399-407) talks about his difficulties with the literary style of H.P. Lovecraft and Kirito (SE01: 226-231) describes often finding rare, thus infrequent, words. Mito (SD01: 304-306) tries to recollect a word he wanted to remember, which was *"urausgefallen"* [really unusual], and Kira (SA02: 182-186) reports frequently coming across technical expressions in *Doctor Who*, which she does not understand despite the use of captions.

While this account is by no means exhaustive, these instances illustrate that students encounter different levels of register as well as different regional varieties in their EE input, although it is interesting to note that participants only referred to inner circle varieties of English. Their statements indicate that EE offers a rich lexical environment in which low-frequency vocabulary or subject-specific expressions can potentially be picked up in addition to providing opportunities for familiarizing oneself with different accents and pronunciations.

In sum, participants in the focus groups corroborate the results of the quantitative strand concerning the most popular EE activities, but at the same time examples of niche activities indicate that engagement with EE very much depends on individualized interests. These results again showcase the great variety of leisure time pursuits involving English among Viennese teenagers in this sample. While the most common EE activities are mainly carried out online

and predominantly involve receptive language use, specialized activities deviate from this tendency and thus expose young Austrians to different types of English. Overall, the amount of exposure to English outside school is enormous in comparison to input at school (see section 4.2.1), as the majority of participants confirm the mean EE time of approximately four hours as plausible. The main reasons for using English for spare time activities seem to lie in a preference for original versions, the greater and earlier availability of content in English, aesthetic qualities attributed to the English language and its use as a lingua franca for communication with family and friends in international contexts.

7.3 Learning from extramural English

One of the main objectives of the focus group interviews was to gain a more detailed understanding of how participants view learning from extramural English. Accordingly, the research questions *How do participants describe learning from EE?* (RQ 5d) and *How do participants evaluate (the potential for) learning from EE?* (RQ 5e) are addressed in this section. Students' descriptions of learning from EE (section 7.3.1) also entail their accounts of what they think they can learn (section 7.3.2) as well as what helps and hinders learning (section 7.3.3). Finally, the learning potential of individual EE activities is analysed (section 7.3.4).

Figure 7.3: Word cloud of the 100 most frequent types in the main category LEARNING FROM EE

7.3.1 Evaluation of learning from extramural English

Most participants believe that they learn from EE as the majority of students across focus groups make positive comments. It is important to clarify here that students were made aware in the interviews that learning did not have to be intentional, but in this case rather meant 'picking something up' while engaging in an activity. The range of what students think they learn varies (see section 7.3.2), but the majority of participants believe that out-of-school engagement with English plays an essential role with 17 out of a total of 30 participants explicitly expressing this opinion and a further five implicitly agreeing with their peers.[202] Several participants argue that

---

[202] Implicit agreement here means that these five students did not verbally respond to the question about learning from EE input, although they might have agreed via gestures such as nodding. However, these students expressed support for the idea of learning in other parts of the interview by making comments on what can be learned or how EE activities help them to remember vocabulary items.

in fact their EE activities contribute more to their knowledge of English than their English instruction at school:

> **Mito**: Ich glaub, das was man in der Schule macht, das das ist ja irgendwie eh ein kleinerer Teil von dem was man von Englisch lernt. Ich mein, wenn (.) ich mein man lernt was, aber irgendwie, allein mit dem was man in der Schule lernt, ich weiß nicht, ich glaub, das (.) also eher ein Großteil lernt man in der Freizeit.
> **Anna**: Ja, das ist wirklich so.
> **SX-f**: Mhm.
> **Anna**: Ich weiß nicht, wenn ich nur in der Schule Englisch mit Englisch zu tun hätte, ich wüsst nicht einmal ein Viertel von dem was ich heute weiß (SD01: 185-188)
>
> **Mito**: I think what we do at school, that that's somehow a smaller part of what you learn from English. I mean, if (.) I mean, you learn something, but somehow, only with what you learn at school, I don't know, I think that (.) well, you rather learn the largest part in your free time.
> **Anna**: Yes, it's really like that.
> **SX-f**: Mhm.
> **Anna**: I don't know, if I only had contact with English at school, I wouldn't even know a quarter of what I know today.

While Mito and Anna agree that they learn some English at school, they emphasize that a large part of their proficiency developed during their free time similar to Marie, who is also of the opinion that she mainly learned English outside school by listening to music, talking to L1 users of English, writing with friends or reading English stories. Marie (SG02: 244-250) even claims that she sometimes already knew language structures by the time they came up in school lessons.

Four participants are not as certain about learning from EE and express mixed views (Lia SA02: 285-293, Maria SC01: 302-306, KingKong SF01: 297-306, Elisa SG02: 220-237) and one further participant (Franz SC01) appears to implicitly agree with this perception. Two of these students, Lia and Maria, seem to think that one could learn from EE, but for them personally that does not happen very often as Lia explains:

> **Lia**: Bei mir ist wenn dann das Vokabular, wegen in den Liedern, aber ich schau keine Filme, also ich schau generell keine Serien oder, ich schau echt nicht viel fern generell und wenn, dann nicht auf Englisch, [...] wenn dann bei Musik mit der Aussprache, aber (1) äh, mit dem Vokabular, weil Aussprache ist ja auch immer so, keine Ahnung, das spricht auch jeder so ein bisschen individuell anders aus und jeder Sänger und jeder Schauspieler finde ich auch spricht's immer so ein bisschen anders aus (2) [...] wenn dann bei den Vokabeln, weil dann (.) oder ich ich erinnere mich wieder an Vokabel wenn ich sie im Lied hör, dann bin ich so 'Ja, das hab ich gelernt, ich weiß was das heißt',
> **Interviewer**: @@
> Lia: oder sowas und dann (.) ja, aber sonst jetzt nicht so wirklich. [...]
> **Lia**: Ich glaub, es würde mir helfen, aber ich mach's halt nicht. (SA02: 285-293)
>
> **Lia**: For me it's vocabulary if anything because of in the songs, but I don't watch movies and well I don't watch series in general or, I really don't watch a lot of TV in general, and if I do, not in English, [...] and if at all then pronunciation through music, but (1) ehm, with vocabulary, because the pronunciation is always so, no idea, everybody pronounces it a little individually and every singer and every actor I think always pronounces it a little different (2) [...] if at all then vocabulary because then (.) or I I remember words again when I hear them in a song, then I'm like 'Yes, I've learned that, I know what it means'
> **Interviewer**: @@
> Lia: or something like that and then (.) yes, but other than that not really. [...]
> **Lia**: It think it would help me, but I just don't do it.

In contrast, Elisa (SG02: 220) and KingKong (SF01: 256) state at first that they do not believe they learn from English at all, but then change their mind in the course of the group discussion. KingKong (SF01: 297-306) explains that he has only been thinking about the type of English he uses to communicate with his friends but has not considered series or other media input. [203] Elisa is unsure about learning from her EE activities for a different reason as her last and concluding statement on this topic shows:

> **Elisa**: Eigentlich find ich, dass mir das auch nichts nützt, weil ich nichts Neues dazulern- also ich benütze Sachen, die ich, also es erfrischt oder hält sozusagen das Niveau auf was ich habe, aber was Neues lern ich nicht dazu meistens. (SG02: 236-237)
>
> **Elisa**: Actually I think that it's not very useful to me because I don't learn anything new, I use things that I, actually it refreshes and it helps to maintain the level that I have, so to say, but usually I don't learn anything new.

This extract suggests that for Elisa learning directly and solely equates to acquiring new knowledge so that the practice effects she evidently describes do not count as learning for her. While no other participant in the focus groups explicitly expresses such an understanding of learning or provides a similar insight into their view of what a learning process entails, this statement highlights that students' conceptualizations of learning may be very different from that of the researcher and thus indicates a need to critically examine notions for which a shared understanding is assumed. While an effort was made to clearly define basic concepts such as 'leisure time' or 'extramural English' in both strands of this study, learning was only construed as including both incidental and intentional processes during the interviews, which could mean that different conceptualizations of learning underlie contributions by different participants.

Finally, only one participant explicitly voices a negative opinion: Lukas is a gamer who argues that most of the international contacts he communicates with while gaming do not speak English well, which leads him to the following conclusion:

> **Lukas**: Nein, xxxxx
> **Interviewer**: @ Warum nicht?
> **Lukas**: Weil ich die meiste Zeit mitm Englischen vorm PC beim Spielen verbringe und die können dann meistens nicht so gut Englisch, da verlern ich's eher. @ (SG02: 232-235)
>
> **Lukas**: No, xxxxx
> **Interviewer**: @ Why not?
> **Lukas**: Because I spend most of the time with English in front of the PC while gaming and they usually do not speak English so well, so it's more likely I unlearn it. @

His classmate Sebastian (SG02: 239-241) supports this statement by adding that he also sometimes feels that his English deteriorates when talking to people who speak English badly, though it is not clear whether he is talking about the context of gaming as well. This view that communicating with others who do not speak English well may hinder learning from EE was expressed several times and will be taken up again in section 7.3.3.

---

[203] It is highly interesting to note that despite the fact that he actively uses English to chat with friends, KingKong does not believe he learns anything from that, implying that in his view the English they communicate in is not good enough to learn from.

7.3.2 What can be learned from extramural English

As discussed in the previous section, a large majority of participants believe that EE activities have a positive effect on their English proficiency. Hence, the topic of this section is what it is that participants think they can learn. Table 7.3 displays all subcategories for WHAT CAN BE LEARNED FROM EE.

Subcategory	Number of coded segments	Number of participants
Vocabulary	12	14
Procedural (rather than declarative) knowledge	7	9
Pronunciation	7	7
More natural English	6	8
Idioms and phrases	5	5
Casual/colloquial English or slang	4	5
Better comprehension skills	4	3
Spelling	3	5
Speaking (freely)	2	2
Grammar	2	2

Table 7.3: Frequencies of occurrence for the subcategories of WHAT CAN BE LEARNED FROM EE

It can easily be seen that vocabulary is most frequently named as an aspect that can be learned from EE, but it is possible that this is due to a research bias: since all interviewees also took part in the quantitative strand, they were aware of the research focus on vocabulary, which could have influenced their responses. However, several other subcategories also point to lexical learning: five students mention that idioms and phrases can be picked up from extramural input and another five refer to casual/colloquial English or slang terms, which also mainly regards lexical expressions. In addition, in line with Nation's (2001) taxonomy both spelling and pronunciation can be considered aspects of word knowledge (see section 3.1.1) as well.

Besides lexical learning, eight participants explain that through their engagement with EE they learn what sounds natural in English because they listen to native speakers in films and series (Johannes SG02: 261-264) or on YouTube (Emma SD01: 345). In addition, they also come across new ways of communicating similar content (Mito SD01: 277). This can give them more confidence in their own language use because if one uses such expressions, one does not run the risk of sounding like a first-year learner of English, as Walküre (SF01: 357) explains. Related to students' impression that they can pick up more natural English expressions, is the idea that they acquire procedural rather than declarative knowledge from EE. This subcategory includes statements about developing a feeling for the language and knowing what sounds 'right' (Paul SC01: 741-745, Susi SD01: 274) without necessary being able to explain why. Most examples provided are about being able to use words and expressions without being able to explain (KingKong SF01: 486) or translate them (Jane SC01: 551-553, Susi SD01: 236-240) or knowing their exact meaning (Elisa SG02: 544-557). Students also point out that the same is true for grammar

with John M. (SD01: 746-748), for instance, explaining that he cannot describe grammar, but he is able to use it because it sounds correct.

Lastly, three participants think that EE helps them to develop their comprehension skills both in the context of reading (Paul SC01: 309-312) and listening (Jane SC01: 317-326, Anna SD01: 351-55). Jane and Anna both attribute their improvement in listening to a habituation effect and explain that over time the effort they initially had to make when listening to spoken English gradually decreased (see also section 7.3.3). Jane (SC01: 334-343) and DJ (SF01: 284-286) are the only two participants that believe EE activities also help with speaking skills: while in DJ's case this belief can be traced back to the large amount of time he spends speaking English during his basketball training (see section 7.2.1), Jane explains that through her father's job in an international organization she has been introduced to a circle of young people from different countries with whom she regularly speaks English. Hence, a tentative conclusion here may be that in participants' perception the development of speaking skills only happens through regular face-to-face contact, however, an improvement in interactive writing skills, which could also be practised over long distances, is not mentioned at all.

7.3.3 What helps and hinders learning from extramural English

In their descriptions of learning participants give insights into their beliefs on how EE activities support their language development and why they think they benefit from them; however, they also mention problems with learning from EE input. This section first outlines participants' thoughts on what helps learning from EE before describing problems identified in the focus groups.

Table 7.4 displays the subcategories for WHAT HELPS LEARNING FROM EE as well as their frequency of occurrence in the data. Nine different ideas on how EE supports language learning are expressed at least once in the focus groups ranging from memory aids over input characteristics and language practices to more ideological assumptions.

Subcategory	Number of coded segments	Number of participants
Input from native speakers	9	10
Repetition	10	9
(Having to) use English actively	8	9
Motivation and own interest	7	7
Familiarization effect	5	4
Using or triggering previous knowledge	4	4
Seeing language in written form	3	3
Being corrected by others	2	3
Linking language to audiovisual content	2	2
Collaborative solving of language problems	1	1

Table 7.4: Frequencies of occurrence for the subcategories of WHAT HELPS LEARNING FROM EE

The belief articulated by the largest number of students is that English is best learned through communication with or exposure to native speakers, especially in terms of pronunciation (DJ: SF01: 389-402, Johannes & Marie SG02: 261-273) and idioms or ways of expressing oneself in English (Mito SD01: 325, Elisa SG02: 380-382). This belief seems to coincide with a notion that language produced by native speakers is more authentic:

> **Johannes**: Ja genau, auch wenn man bei Filmdialogen zuhört, kann man ein bisserl sehen wie die wie die reden, die
> **Interviewer**: Ok.
> **Johannes**: also die echten Englischsprachigen und das kann man sich dann auch ein bisschen merken.
> **Interviewer**: Ja.
> **SX-m**: <ENGLISH> native speaker. </ENGLISH>
> **Johannes**: <ENGLISH> native speaker. </ENGLISH> (SG02: 261-267)

> **Johannes**: Yes exactly, if you listen to dialogues in films, you see a little how they how they speak the
> **Interviewer**: Ok.
> **Johannes**: well the real English speakers and then you can also memorize that a bit
> **Interviewer**: Yes.
> **SX-m**: <ENGLISH> native speaker. </ENGLISH>
> **Johannes**: <ENGLISH> native speaker. </ENGLISH>

Native speaker input is also regarded as being of a higher quality; this was mentioned several times by contrasting non-native speakers who may make mistakes (see below) with L1 speakers. Accordingly, some students attribute more importance to receiving L1 English input because, as Mito (SG02: 325) puts it, in order to learn English, one needs to have a source of 'good English'. This relatively strong unquestioned belief is interesting and slightly troubling from a sociolinguistic perspective, which will be further discussed in section 8.2.3.

In a few instances, reference to native speakers is connected to the idea that using or even being forced to use English productively is more beneficial than receptive use. Elisa (SG02: 380-382) and Keanu (SA02: 294-296) are convinced that speaking to natives is more helpful than other activities and Louise (SG02: 356-365) thinks that English or any other language is best learned by travelling to countries where it is spoken. Others express the notion that productive use is more beneficial without linking it to L1 English users: DJ (SF01: 2-10) refers to his basketball training, Marie (SG02: 244-250) thinks that speaking or chatting to friends helps her more than reading, Susi (SD01: 314-315) provides the example of having to reflect on language while formulating messages, and John W. (SA02: 305) reports his experience of using English as a lingua franca during a diving course in South America.

Another characteristic of EE input that is seen as very helpful for learning is repetition, both in terms of repeated encounters and of repeatedly engaging in an activity, like for instance listening to the same song (John SE01: 372-374, Walküre SF01: 549-553) or playing the same game more than once. Several students assert that coming across words and expressions repeatedly during EE activities helps them to understand their meaning and learn them incidentally (Maria & Franz M. SC01: 302-306, Jane SC01: 376, Mito SD01: 227-230, Walküre: SF01: 479-485), both from written and spoken input. Examples given are the 'catchphrases' in songs (Kirito SE01: 370),

recurring expressions in series (Karl SE01: 249) and computer games (Elisa SG02: 453), or repetition through the linguistic habits of YouTubers (Anna SD01: 353-355).

Linked to repetition is the idea that EE helps language development because it triggers previous knowledge or because students have to use already existing knowledge in their extramural activities. Elisa's comment quoted in section 7.3.1, in which she argues that she does not learn new things, but that EE helps to maintain her level of English, is a prime example for students' belief that recycling language structures outside school helps to consolidate their language competence. Kira (SA02: 409-414) and Louise (SG02: 643-645) give examples of how encountering English words in different contexts both inside and outside of the classroom caught their attention and thus helped them to remember these lexical items.

Repetition is also at the basis of the habituation effect that already came up in the previous section with Anna, Jane and Paul reporting that over time they found listening or reading increasingly easier. Jane describes an example of involuntarily listening to conversations in English in the street or on public transport to explain how much easier comprehension of spoken English has become for her and concludes:

> **Jane**: Jaja, ich konzentrier mich automatisch auf irgendeine auf was Ungewöhnliches, in unserem Fall ist das Englisch, ahm, und ja mit der Zeit ist es dann ganz gewöhnlich geworden und ich muss mich überhaupt nicht anstrengen, ich hör einfach, also ich muss nicht einmal zuhören, ich hör trotzdem. Weil früher musste ich so ‚Ok, was was sagen die jetzt?',
> **Interviewer**: Mhm.
> **Jane**: jetzt steh ich einfach und ich äh, also ich bekomme, äh es kommt trotzdem (SG02: 322-324)

> **Jane**: Yeah, yeah, I automatically concentrate on anything on something unusual, in our case that's English, ehm, and yes with time it has become really normal and I don't have to make an effort at all, I just listen, well I don't even have to listen, I hear it anyway. Because before I had to be like 'Okay, what do they say now?'
> **Interviewer**: Mhm.
> **Jane**: now I just stand there and I eh, well I just get eh it comes anyway

Such a familiarization effect could potentially also enhance enjoyment of EE activities, since a decrease in effort needed for comprehension makes watching an English series or reading in English more similar to watching or reading for pleasure in the L1.

In addition, a belief that EE activities support language learning because they are done out of genuine interest and thus enhance motivation is articulated several times, especially in contrast to lessons at school. In describing past EE activities, DJ and in particularly Kirito argue that it was a specific interest that first got them involved in extramural activities and then increased their overall motivation. The fact that EE activities are their own choice is generally evaluated extremely positively (Lia SA02: 422) and students argue that because they are really curious about the content, they are also more interested in the language expressing it (Walküre SF01: 291-293, Franz, Pinguin & Kirito SE01: 450-465, Marie SG02: 489-493).

Furthermore, students state that EE activities help them to memorize language structures in different ways. Some find it helpful to see language in written form through subtitles (Keanu

SA02: 176-179) or in books (Jane SC01: 374-376, Susi SD01: 312-315), while others think that linking language structures to audiovisual content works well as an aid to memory (Vanessa SD01: 336-342, Marie SG02: 390-397). Moreover, EE activities can sometimes involve an explicit focus on language, as examples from online contexts provided in group SF01 suggest. DJ, KingKong and Walküre (SF01: 359-386) describe instances of wrong language being corrected on social media platforms:

> **KingKong**: Ich glaub bei den <ENGLISH> social media </ENGLISH> Plattformen kann man auch ganz viel lernen. Weil ist ja so halt international ganz viele Leute aus ganz vielen Ländern, und es ist halt ganz lustig, wenn einer einen Fehler macht, dann wird er in den Kommentaren immer meistens <ENGLISH> totgehated </ENGLISH>,
> **Interviewer**: Echt?
> **DJ**: Jaja.
> **KingKong**: wegen dem Fehler. Das wird dann immer sofort drauf hingewiesen und <6> da kann man auch ein bissi was lernen. </6>
> **Walküre**: <6> Mh, da gehört ein anderes Wort hin </6> (SF01: 359-363)

> **KingKong**: I think on <ENGLISH> social media </ENGLISH> platforms one can also learn a lot. Because it is so international a lot of people from very many different countries, and it's quite funny, if someone makes a mistakes, then he usually is <ENGLISH> hated to death </ENGLISH> most of the time in the comments.
> **Interviewer**: Really?
> **DJ**: Yeah, yeah.
> **KingKong**: because of the mistake. It's then always immediately pointed out and <6> from that you can also learn a little bit </6>
> **Walküre**: <6> Mh, there needs to be a different word there </6>

The three students explain that such instances of 'public shaming' and correction in social media discourse often concern mistakes such as wrong word choice or mixing up *their* and *they're* (SF01: 363-386) and that although it's often not done nicely, mistakes are usually corrected. In addition, DJ (SF01: 258-262) also provides examples from his basketball team (see section 7.2.1), where in the beginning he was corrected by his teammates if he said something wrong, though in a friendly way. Finally, Walküre (SF01: 404-409) also provides a more positive example of language-focused learning in an online context by describing how international members of an online group she participates in sometimes collaboratively solve language problems. She reports that when they discuss a specific topic people may not know certain words in that area and the suggestions supplied by different community members then present a nice learning opportunity. It is important to note that the communities these students describe include people from all over the world, the majority of which are not likely to be L1 speakers of English. These experiences therefore constitute examples of learning from ELF, which stand in contrast to more frequently voiced belief that English is best learned from native speakers.

In addition to the many aspects that participants believe help language learning, they identify a number of problems that hinder learning from EE. Table 7.5 presents the subcategories of PROBLEMS WITH LEARNING FROM EE together with their frequency of occurrence.

Subcategory	Number of coded segments	Number of participants
Learning wrong things from people who speak 'bad English'	7	7
In songs language is different	5	6
Learning incorrect grammar or spelling	3	3
Making wrong inferences	2	2
English sometimes too difficult	1	2
Not encountering language structure often enough	1	1
Not learning anything new due to repetition	1	1

Table 7.5: Frequencies of occurrence for the subcategories of PROBLEMS WITH LEARNING FROM EE

The most frequently named problem can be considered the reverse of the belief that English is best learned from native speakers as seven students express a concern about learning incorrect language from people who speak 'bad English'. Lukas and Sebastian (SG02: 234-241, 408-413), who do not think that they learn from EE input (see section 7.3.1), represent this view by referring to gaming and commenting on the poor English proficiency of other gamers. Similar concerns are voiced by Kirito (SE01: 340), who argues that gamers on the European server do not pay attention to grammar, spelling or capitalization with 'catastrophic' results. Such a situation could then lead to teaching each other wrong structures in Franz's opinion (SE01: 332-334), a view that is echoed in relation to other types of (online) communication among non-natives by Mito (SD01: 321-325), DJ (SF01: 402) and KingKong (SF01: 297-302).

Learning incorrect grammar or spelling is however also mentioned without reference to interactive communication. Karl (SE01: 326-331) thinks a lot of language found on social media is incorrect and both KingKong (SF01: 612-623) and Marie (SG02: 370-372) mention picking up wrong grammatical structures from songs: KingKong recalls a rap line saying "I wish you was never born", which he thought was accurate until his teacher corrected him, and one of Marie's favourite songs includes the line "he don't adore ya". Interestingly, although she is well aware that this structure is considered grammatically incorrect in standard English, Marie thinks that she takes up such wordings from lyrics subconsciously and explains that as a result she has to take care to correct herself in her own language use, particularly at school. In fact, songs were the only specific EE activity mentioned several times as being problematic to learn from because the pronunciation is often different when sung (Lia SA02: 287-289, Mito SD01: 319-321) and word order may diverge from standard speech to fit the melody (Jane & Maria SC01: 657-663). Besides, the lyrics are often hard to understand (Franz SE01: 365) and frequently include slang terms that one cannot use at school (Karl SE01: 375).

In addition, Marie (SG02: 542) and Susi (SD01: 466-469) note a danger of making wrong inferences from EE input, for instance with regard to the meaning of unknown words, and Kira and Lia (SA02: 182-188) argue that sometimes the level of English used in series is too difficult for them to pick anything up, particularly if technical terms relating to science or medicine are used. Finally, the last two issues with EE input identified by participants are two sides of the same coin: Elisa (SG02: 435-458) complains that nothing new is learned from EE because of

repetition and Paul (SC01: 540) states that language structures are not encountered often enough to actually learn them. These two statements clearly refer to problems related to frequency (see section 3.1.3): while frequent language structures form a large part of everyday language use and are recycled again and again, less frequent structures are not encountered often enough to actually acquire them. Such frequency effects are, however, present in all situations of language use and learning and not specific to EE.

7.3.4 Evaluation of the learning potential of different EE activities

In the focus groups, participants were presented with a second visual input depicting many different leisure time pursuits (see Appendix B) to stimulate a discussion on whether some EE activities were more helpful for language learning than others. This section presents students' evaluations of the learning potential of different activities, many of which are interconnected with the beliefs outlined in the previous section.

Using English productively in face-to-face conversations or in online communication is probably considered the most beneficial EE activity (e.g. Jane SC01: 334-344, Maria & Franz. M SC01: 361-366). Marie (SG02: 372) and Elisa (SG02: 380-382) argue that speaking is even more effective than writing because in chats the language is usually very simple and many abbreviations are used, whereas spoken interaction can be more complex and it is easier to ask for explanations. Others like Anna (SD01: 327-332) consider chatting and written online communication very useful and Pinguin (SE01: 324-326) also evaluates social media positively. His classmate Karl (SE01: 323-331) disagrees because he thinks there are too many mistakes in social media discourse; these, however, are construed as a learning opportunity by KingKong, who argues that one can learn from social media precisely because mistakes are often corrected by other users (SF01: 359-386, see section 7.3.3). These examples indicate that based on their own experiences and beliefs participants' perceptions of learning opportunities differ widely and that evaluations of learning potentials can only be summarized at a very general level.

In addition to using language, reading was seen as helpful for learning by a number of students (Jane SC01: 368-376, Marie SG02: 483-491). Susi (SD01: 312-315) for instance argues that seeing language in written form is best for her because she prefers a visual learning style and Walküre (SE01: 356-357) thinks reading novels definitely helps to attain a higher level of English and to develop a larger vocabulary because one encounters many synonyms. Another reason why reading is considered more beneficial than other activities is provided by Jane and Paul in the following extract:

> **Jane**: Beim Lesen ist man mehr konzentriert als beim Filme schauen.
> **Paul**: Genau, das glaub ich nämlich auch, dass man sich beim Lesen mehr drauf konzentrieren muss und,
> **Jane**: auf die Wörter konzentriert,
> **Paul**: genau, und genau auf den Wortlaut und sowas konzentrieren muss als bei Serien, weil wenn ich in einem Buch zum Beispiel ein Wort seh, dass ich nicht verstehe, dann bleib ich wahrscheinlich stehen beim Lesen und les nicht weiter drüber, bei einer Serie, die Serie läuft weiter, da werd ich sagen ‚Ja, ok' (SC01: 598-602)

> **Jane**: While reading you concentrate more than while watching films.
> **Paul**: Exactly, I think so too, that you have to concentrate more on it while reading and
> **Jane**: concentrate on the words
> **Paul**: exactly, and concentrate exactly on the wording and such, you have to concentrate more than for series because if I see a word I don't know in a book for instance, then I probably stop reading and do not read over it, but in a series, the series just continues, then I say, 'Yes, ok.'

Paul and Jane argue that reading requires more concentration on the language overall and they add that in contrast to series they can control the processing speed and therefore they are more likely to stop at unknown words and take time to think about them. At a different point in the interview Paul (SC01: 378-380) emphasizes that reading in English is more difficult compared to watching films or series because there are no visuals to aid comprehension and thus unknown language has a greater impact and makes reading more taxing, which suggests that while reading may lead to greater learning gains in his opinion there can also be greater inhibitions to engage in this more demanding activity. This aspect is also taken up by Maria (SC01: 369-370) who has a more mixed view on the benefits of reading because it requires effort and because some teenagers do not even like to read in German.

Audiovisual input like films, series and YouTube videos was considered best for learning by fewer participants, but for Keanu (SA02: 290-293), Vanessa and Emma (SD01: 336-347) it is the activity of their choice. In the following extract Vanessa explains why she thinks audiovisual input supports her learning:

> **Vanessa**: Naja, bei mir ist das so, wenn ich Serien oder Filme anschau, lern lern ich am besten, weil da ist die Handlung vor dir und du hörst dann auch die Aussprache und,
> **Interviewer**: Aha, ja.
> **Vanessa**: und das kombiniert sich dann so. (SD01: 336-339)

> **Vanessa**: Well, for me it's like that, when I watch series or films, I learn learn best, because there the plot is in front of you and you also hear the pronunciation and
> **Interviewer**: Uh-hu, yes.
> **Vanessa**: and then these combine in a way.

Vanessa's classmate Emma agrees that the combination of the visual image and spoken input is a great support for both comprehension and learning.

The question whether gaming supports English learning led to rather animated discussions, particularly in groups SC01 and SG02. The general conclusion was that 'it depends' and the main aspect learning depends on is the type of game in participants' opinion. Students did not agree whether one can learn from single-player games in contrast to online multiplayer games, which are generally not seen as helpful. In the data there is a great range of positions adopted on this issue: some students think that video games are great for language acquisition in general because you have to understand the language in order to know what to do (DJ, KingKong & Walküre SE01: 312-340), while others argue for benefits of single-player games over multiplayer games (John M. SC01 408-457, Louise SG02: 398-481) or the other way round (Maria SC01 408-457). Moreover, frequent gamers such as Lukas (SG02 398-481) or Kirito (SE01: 332-346) are more sceptical when it comes to learning benefits of games.

Lukas explains that he mostly plays multiplayer games online and does not really pay attention to "what the game babbles in his ear" (SG02: 406). In addition, many gamers on the server he

plays *Counterstrike Global Offensive* on are from Russia and do not speak English well, which leads him to the already mentioned conclusion that he is more likely to unlearn English while speaking or chatting with them (see section 7.3.1). In contrast, Louise expresses a strong belief that games are very helpful for learning English based on her experience with single-player games. She argues that the input from the game itself is definitely correct and gives the example of learning from educational games when she was at primary school level.[204] Other participants enter this discussion on one side or the other resulting in a highly passionate dispute as exemplified by the following extract:

> **Marie**: Genau, weil ich, keine Ahnung, so solche Spiele wie, keine Ahnung, COD oder so, so auf die Art da kann man mit, man kann mit Leuten (.) man hat man liest da die ganze Zeit was auf Englisch und es tut schon irgendwie das English verbessern halt einfach, es kommt auch halt drauf an wie (manche) schon gesagt, auf das Spiel.
> **Interviewer**: Ja. Klar.
> **Lukas**: Ja und Onlinespielen bezieht sich ja eigentlich das was vom Spiel kommt, das es wo das Englisch sicher richtig ist meistens nur so etwas wie <ENGLISH> an enemy has been slayed </ENGLISH> auf solche Sachen halt einfach.
> **SX-f**: <8>xxxxxxx </8>
> **Marie**: <8> <LOUD> Naja es kommt manchmal auch was über Missionen, tu das und das </LOUD> </8> (SG02: 424-428)

> **Marie**: Exactly, because I, no idea, such as games as, no idea, COD [*Call of Duty*] or so, well somehow you can with, you can with people (.) you have you read something in English the whole time and that simply makes your English improve, it just depends on the game as (others) have already said, on the game
> **Interviewer**: Yes, of course.
> **Lukas**: Yes and in online games what comes from the game, where English is certainly correct, usually only refers to things like <ENGLISH> an enemy has been slayed </ENGLISH>, simply on things like that.
> **SX-f**: <8>xxxxxxx </8>
> **Marie**: <8> <LOUD> But sometimes there is also something about missions, do that and do that </LOUD> </8>

Marie, who quite clearly is not a gamer herself, tries to make the point that during gaming one does not just get in contact with the language produced by other players, but also with the input from the game itself, which Lukas refutes explaining that the same simple phrases are repeated over and over again. It can be argued that the phrase "an enemy has been slayed [sic]" is not that simple from a language acquisition point of view and could actually be a prime example of vocabulary learning from EE input. However, Lukas clearly would not agree with such a view and the fact that he produces an incorrect past participle for the verb *slay* may actually be testament to his statement that he generally does not pay attention to input from the game itself.[205]

What emerges from the discussion in group SG02 is that in students' opinion much depends on the type of game played and this line of thought is further explored in group SD01. After a similar disagreement on the benefits and drawbacks of multiplayer and single-player games, Paul argues that it actually depends more on how story-based a particular game is:

---

[204] From the interview data it is not completely clear whether Louise still plays games nowadays, she does not position herself as a gamer in the same way as Lukas.
[205] As Lukas provides this example several time in the interview, it is clearly discernible that he produces *slayed* rather than *slain*.

> **Paul**: Also es kommt wirklich sehr drauf an, also wie wie wie storybasiert das Spiel ist, also ja es kommt schon darauf,
> **John M.**: Es muss nicht mal (xx) sein, die Story,
> **Paul**: Jaja, <19> es kann es kann verschiedenes genau </19>
> **John M.**: <19> es gibt ja auch verschiedenes Zeug </19> was du auf Englisch lesen kannst,
> **Paul**: Genau. Also, wenn du vielleicht ein langes Spiel alleine spielst, halt auf Englisch, mit vielen Dialogen und sowas, kannst du wahrscheinlich auch, so wie von Serien, viel, sehr viel mitnehmen. Wenn du online spielst mit anderen Leuten, kann man wahrscheinlich mehr vom Selberreden mehr mitnehmen, ahm, wohl wahr ist es halt oft, dass es vielleicht nicht grad das perfekteste Englisch ist. (SC01: 425-429)

> **Paul**: Well, it really depends a lot on, well on how how how story-based the game is, well yes it just depends on
> **John M.**: It doesn't even need to be (xx), the story,
> **Paul**: Yeah, yeah, <19> it can it can be different things exactly </19>
> **John M.**: <19> like there also is a lot of different stuff </19> you can read in English
> **Paul**: Exactly. So, if you play a long game on your own, well in English, with a lot of dialogue and such, you can probably, like from series, pick up a lot, really a lot. If you play online with other people, you can probably take away more from speaking yourself, ehm, it's true that often, that it's not really the most perfect English.

It is interesting to note that Paul compares single-player games to series, a statement which is reinforced later on when the students explain that like for TV series and films, the original language version also tends to be better for games (SC02: 483-485, see section 7.2.3). In sum, gaming was the subject of the most intense discussions concerning the learning potential of EE activities, and opinions on whether it is beneficial for language development and, if so in what ways, range from negative effects over the acquisition of simple conversational phrases, slang terms and abbreviations to fervent support.

While the debates on learning from music were not quite as lively, evaluations also vary greatly as, in fact, music is the only activity whose learning potential is evaluated negatively by six participants. Some problems identified by participants like varying pronunciation or non-standard grammatical structures have already been discussed in the previous section, but another issue that emerges several times is a lack of concentration on the lyrics because the melody is in the foreground (e.g. Jane & Maria SC01: 351-355). Franz makes a similar point when she states that there is a big difference between listening to music and concentrating on it or listening to music as an accompaniment while doing something else:

> **Franz**: Also ich glaub man kann schon [lernen], aber man muss sich halt wirklich voll da drauf konzentrieren, weil, also wenn ich einfach so nebenbei englische Musik hör, muss ich ehrlich sagen, ich würde gar nichts verstehen. (SE01: 365)

> **Franz**: Well I think you can [learn], but you really have to fully concentrate on it, because, well, if I simply listen to English music on the side, I honestly have to say, I would not understand anything.

For these reasons, participants like Anna (SD01: 177-181) and Mito (SD01: 321) do not believe that language and especially vocabulary is learned from songs, whereas Lia (SA02: 285-287) thinks if she learns English outside school then it is definitely from music. Kirito (SE01: 370) argues that some recurring parts like the catchphrase are learned without even paying attention and his classmate John (SE01: 372-374) also evaluates the learning potential of music as fairly positive because he listens to the songs over and over again, which helps him understand the

lyrics and also leads to learning. Similarly, Walküre (SF01: 549-561) explains that the more she listens to songs, the more she learns the lyrics by heart and then she is really interested in their meaning and takes the time to look them up.

Consequently, in addition to repetition effects (see section 7.3.3) paying attention to and/or looking up the lyrics emerges as a key feature of learning from English songs in several participants' statements. Paul and Jane (SC01: 609-611) explicitly agree with this view, but later Jane and Maria add that many popular songs do not have proper texts:

> **Jane**: Ja, wenn man in einem Lied so meistens <ENGLISH> Oh no, oh yeah, yeah, yeah </ENGLISH> dann ist das,
> **Interviewer**: Ja, stimmt. @ Das ist auch nicht mehr wirklich,
> **Maria**: Manchmal hat ein Lied nicht mal nen Text. Das ist dann immer so huiiii,
> **Jane**: Oder dieses lalalalala @@ (SC01: 632-635)
>
> **Jane**: Yes, if in a song you mainly have ENGLISH> Oh no, oh yeah, yeah, yeah </ENGLISH> then that's,
> **Interviewer**: Yeah, right. @ Then that's also not really
> **Maria**: Sometimes a song doesn't even have a text. That's then always like huiiii
> **Jane**: or this lalalalala @@

The same opinion that lyrics frequently do not even make sense is voiced by Pinguin (SE01: 399). The students in his group (SE01: 364-423) agree that the learning potential of listening to music depends not only on whether one concentrates on the text, but also on the music genre. Moreover, they concur that singing, for instance karaoke, is probably more beneficial for language learning than only listening to music. Taking up this group's realization that genre is important, it is interesting to note that participants in two other groups (SD01 and SF01) come to the conclusion that hip hop may be particularly good for learning English, especially old-school rap, because there is a clear focus on the lyrics, the texts include topics like racism that inspire reflection (SF01: 498-547) and the language used in rap songs is often quite complicated (SD01 637-654).

In sum, findings presented in this section show that a large majority of participants believe that their EE activities have a positive effect on their English proficiency. The students identify a wide range of aspects that can be learned from EE: vocabulary and other lexical features are named most frequently followed by pronunciation and procedural knowledge of English which leads to more natural language use. Participants exhibit diverse beliefs on what helps or hinders learning from out-of-school English input. Many appear to be strongly influenced by a conviction that English is best learned from L1 speakers and that input from non-natives is potentially problematic. In addition, repeated encounters and a familiarization effect are regarded as helpful, as is active language use. Participants also emphasize that EE activities have a positive impact on motivation because they engage in them of their own volition. In total, fewer problems than advantages are named by the focus group members, which again points to an overall positive perception of extramural English by the students in the focus groups. In the discussions on the learning potential of different EE activities productive language use through speaking and chatting, reading, and, to an extent, viewing audiovisual media are evaluated as very helpful for language learning, whereas the learning potential of music and even more so gaming was the subject of heated debates and is assessed very differently by different participants.

7.4   Vocabulary learning from extramural English

Like the previous section, this one is concerned with learning from extramural English, but this time focusing on vocabulary as the main interest of this study. The second part of the answer to the research question *How do participants interpret and evaluate the results from the quantitative part of the study?* (RQ 5c) is provided first by describing participants' reactions to the results on vocabulary learning strategies. The other question addressed in this section, *How do participants describe their practices of vocabulary learning from EE?* (RQ 5f), highlights students' accounts of lexical learning. In their discussion of vocabulary acquisition participants mainly provide information on the following aspects: their use of strategies to discover the meaning of unknown words and which words are actually looked up (section 7.4.1) and in how far they remember these new words or what helps them to memorize their meanings (section 7.4.2). In addition, the last section also includes examples of lexical items learned from EE input.

Figure 7.4: Word cloud of the 100 most frequent types in the category Vocabulary learning from *EE*

7.4.1 Using strategies to discover the meanings of new words

The third and final visual stimulus participants were given during the focus groups was a bar chart summarizing the quantitative results on the use of vocabulary learning strategies obtained from the EEQ (see Appendix B). As described in section 6.1.1, the most popular strategy to deal with unknown words encountered during participants' spare time activities is guessing from context, both in the case of audiovisual and print media, followed by looking up words in (online) dictionaries. Among the language-based strategies only comparisons to other languages are regularly used by a majority of students, whereas few consider the part of speech or attempt to analyse word parts. Finally, the social strategy of asking someone for help is not often used either. However, overall almost 60% of the students in the quantitative strand claim that they usually do something when encountering a new word.

Overall, participants in the focus groups support the quantitative findings on discovery strategies. In the descriptions of participants' own practices, guessing from context clearly emerges as the default option (John W. SA02: 397, Paul SC01: 581-590, Emma SD01: 449, DJ &

KingKong SF01: 454-455). This strategy is used in combination with all kinds of input and is regarded as especially useful:

> **John**: Bei mir ist das oft so, dass ich versuch, also die Wörter, die ich nicht versteh, quasi im Zusammenhang zu verstehen was das bedeuten könnte, ich mach das halt so,
> **Interviewer**: Mhm.
> **John**: ich googles halt nicht direkt nach, weil das sind halt schon einige Wörter, die man dann irgendwie nicht versteht. (SE02: 263-265)

> **John**: For me it's often the case that I try to, well the words I don't understand, to understand them within the context what that could mean, I just do it like that
> **Interviewer**: Mhm.
> **John**: I don't google it directly because after all it's quite a few words you somehow don't understand.

John implies that inferring the meaning of unknown words from context takes less effort than looking them up. Similarly, some students (Marie, Louise and Elisa SG02: 579-584) argue that guessing happens automatically and thus they find it questionable whether inferring from context even counts as a strategy, or as they put it, as 'doing something':

> **Marie**: Du liest so einen Text und dann kommt so ein Wort wo ‚Achso, ok,' ich werd nicht mal so, ich find das passiert irgendwie voll natürlich, auf einmal so voll normal so irgendwie man sieht ein Wort, man weiß nicht was es ist, normale Reaktion denk kurz drüber nach, kenn ich, kenn ich's nicht, interessiert's mich, schau weiter. (SG02: 584)

> **Marie**: You just read a text and then comes such a word where 'Ah, okay', I don't even, I think that somehow happens completely naturally, at once totally normal, kind of you see a word, you don't know what it is, normal reaction, quickly think about it, do I know it, do I not know it, does it interest me, go on.

While this process may be more or less automatic for different students, it is the preferred method of handling unknown words across groups and is particularly adequate for some of the most frequent activities. When watching series, films or videos students only have very little time to find out about the meaning of a new word. Hence, strategies like asking or using a dictionary take too much time and would require the programme to be stopped, whereas inferring from context is regarded as a more viable option in such cases (SA02: 402-408). The only problem identified with guessing from context is that one only knows what a new word roughly means and cannot use it (Susi SD01: 466-469), but arguably in many contexts approximate comprehension is sufficient for understanding the content, especially when engaging in EE activities for entertainment purposes.

While guessing from context is positioned as the default option across groups, some participants also report thinking about other languages and using dictionaries. Comparing a new word to other languages in their repertoire is preferred by a number of students (e.g. Susi SD01: 411, 426-428, DJ SF01: 432-434, Elisa, Johannes, Louise & Marie SG02: 505-512) with most implying that only if that does not resolve the problem, they might go on to use further resources like dictionaries. Dictionaries are frequently discussed in the focus groups, usually in form of online platforms and apps. Looking up a word, particularly on a smartphone, is described as easy and fast (Maria SC01: 511-513, Susi SD01: 407, Walküre SF01: 436, Sebastian SG02: 517-519), but for most it is still the last resort (Paul SC01: 540-544). As suggested in the quote below, dictionaries are only used if all other strategies fail:

> **Anna**: also wenn ich damit klar komm, dann komm ich damit klar und wenn's wirklich so irgendwie das Hauptwort des gesamten Textes ist, dann (2) dann überleg ich halt was ich vorher alles machen kann, bevor ich nachschaue, weil ich zu faul dafür bin @ ein englisches Wort nachzuschauen, weil das ist mir auch zu blöd dann. Aber wenn ich's dann wirklich nicht weiß, dann so als letzte Möglichkeit schau ich dann nach. (SD01: 429-432)
>
> **Anna**: well, if I can manage, then I can manage and if it's really like somehow the main word of the whole text, then (2) then I think about everything I can do before looking it up, because I am too lazy for that @ to look up an English word, because that's stupid and too much of an effort. But if I then really don't know, then as the last possibility, then I look it up.

In contrast to this view, some participants actually favour dictionary use: Walküre and KingKong (SF01: 438-441), for instance, agree that thinking about word meaning is more strenuous than looking a word up. Kirito (SE01: 123) states that using dictionaries is his preferred strategy because it gives you the exact word meaning and for this reason dictionaries are sometimes also used to check inferences based on guessing from context (Paul SC01: 545). Pinguin (SE01: 425-427) as well as Maria (SC01: 511-513) point out that using an online dictionary is definitely much quicker than asking someone:

> **Maria**: Ich mach das immer so. Ich schau was, les was oder irgendwas, dann nehm ich mein Handy raus und gib das Wort Wort ein,
> **Paul**: Ja es, das geht wirklich schnell.
> **Maria**: weil das ist einfach am schnellsten. Weil selten hat man dann irgendwen neben sich, der das, wovon man weiß, dass der das Wort auch kann oder dass er einen größeren Sprachgebrauch hat als man selber, also Wortschatz. (SC01: 511-513)
>
> **Maria**: I always do it like that. I watch something, read something or whatever, then I take out my phone and type the word in
> **Paul**: Yeah it, that's really fast.
> **Maria**: because that's just the quickest. Because rarely there is someone beside you who, of whom you know that he knows that word or that he has a larger language use than yourself, well vocabulary.

Interestingly, asking someone for the meaning of new English words is a relatively unpopular strategy. While a few students state that they like being able to ask others (e.g. Franz SE01: 267-269) and sometimes do so (Emma SD01: 449), most participants present arguments against the use of such a social VLS. First, of all someone needs to be in the vicinity, but that is often not the case (Johannes SG02: 524-528) because students do many EE activities on their own (Paul SC01: 490). In such situations one could then contact others via chat, but then it takes time until one gets an answer (Pinguin SE01: 425-427). Second, the person one wants to ask needs to be more proficient in English than oneself (Elisa SG02: 523) and be able to explain the word (Kirito SE01: 429-433). This last point is also stressed by Lia:

> **Lia**: Was mir dazu einfällt ist, ich glaub jemanden fragen ist einfach zu wenig, ich mein das soll jetzt nicht arrogant klingen, aber (1) wenn wir das Wort nicht wissen, wieso sollt das dann jemand in unserer Um-, wenn wenn der jetzt kein nicht in England gelebt hat oder, (SA02: 389)
>
> **Lia**: What comes to mind for me is, I think asking someone is just not enough, I mean, that's not supposed to sound arrogant now, but (1) if we don't know the word, why should someone know it in our en-, if if that person hasn't lived in England or,

Furthermore, in addition to needing someone who is able to answer questions on word meaning, one also needs someone who is willing to do so. This may not always be the case with siblings, as Mito (SD01: 462) explains that he gets really annoyed at his younger brother for asking him

about English words because he could as easily look them up on his phone. Finally, some students also report inhibitions to ask others because it could be embarrassing, for instance if it is a really easy word one just cannot think of in that very moment (Paul & Maria SC01: 517-520, Susi SD01: 407).

The remaining strategies of identifying the part of speech and analysing word parts are rarely employed because they are seen as taking too much time (Franz SE01: 440) and requiring too much effort (Walküre SF01: 422-431). In addition, students argue that not every English word can be separated into parts (Lia SA02: 408) or at least that it does not work as well in English as in other languages, like for instance in Latin (John M. & Jane SC01: 498-500).

From the interview data it thus becomes clear that guessing from context is the default option for most participants and that only if this strategy does not help to resolve the issue, the students resort to other VLS such as thinking about other languages and using online dictionaries or apps. Looking words up in dictionaries is seen as fast and easy, but still only used if the meaning of an unknown word cannot be readily inferred from context. This finding raises the question what makes a new word encountered in EE input important enough to look it up or employ additional VLS.

A first characteristic that can be inferred from the discussion of guessing from context above (see Anna's quote) is that new lexical items are only worth the effort of looking them up if they are needed to understand the content *and* cannot be easily inferred from context. If new words are deemed unimportant, participants simply ignore them (Lia SA02: 398, Walküre SE01: 342). Irrelevant stretches of speech or text, described by Elisa (SG02: 564) as "random talk", do not warrant learners' strategic attention as explained by Anna:

> **Interviewer**: Mhm. Aber eben (4) wie hast du jetzt so schön gsagt, wenn ich wenn ich damit klarkomm, heißt ja, heißt das dann, dass du, wenn's dir grad nicht einfällt oder so, aber dass du's aus dem Zusammenhang, dass du den Sinn ungefähr verstehst, oder?
> **Anna**: Nicht mal, teil- teilweise nicht mal das, sondern einfach nur wenn es einfach ein unwichtiges Wort ist,
> **Interviewer**: Ahso,
> **Anna**: das ignorier ich einfach komplett. (SD01: 433-436)

> **Interviewer**: Mhm. But just now (4) how did you put it so nicely, if I if I can manage, that means, does that mean that you, that you can't think of it in that moment or so, but that you get it from context, that you roughly understand the sense?
> **Anna**: Not even, some- sometimes not even that, but only if it simply is an unimportant word,
> **Interviewer**: I see,
> **Anna**: then I completely ignore it.

Hence, the perceived importance of a word appears to depend, firstly, on how essential it is for being able to follow the content and, secondly, on how much information is provided by co-text and context. Still, this is not the only characteristic that makes looking up lexical items worthwhile. Two other criteria found in the data are repeated encounters with the same word and arousing students' interest. The idea that words which are encountered repeatedly merit further investigation is expressed by Jane and Maria in group SC01 (529-531) and Mito in group

SD01. Mito seems to have reflected on his practices relating to unknown vocabulary during the interview and makes this statement after the discussion has been going on for a while:

> **Mito**: Ich mach das eigentlich, fallt mir jetzt ein, meistens so, wenn ich irgendwie wenn ich ein Wort mal so sehe, dann les ich meistens drüber, aber wenn ich irgendwie so jetzt bei gewissen Wörtern wenn mir da einfällt da hab ich jetzt eigentlich schon ziemlich oft drübergelesen, ahm,
> **Interviewer**: Aha.
> **Mito**: dann denk ich mir ‚Es ist jetzt mal langsam an der Zeit, das nachzuschauen, weil sich's lohnt'.
> **Interviewer**: Ok.
> **Mito**: Ich mein, ich schau eigentlich eh ziemlich viel nach, weil ich ja immer wieder neue Wörter krieg, ah, seh, wo ich mir denk ‚Ah, das hab ich mich jetzt eigentlich schon oft gefragt x dass ich das nicht kenn', ja. (SD01: 454-458)

> **Mito**: I actually do it, I now realize, usually like this, if I somehow if I just see a word, then I usually skip over it while reading, but if I somehow with certain words if I realize that I have actually skipped over them quite often lately, ehm
> **Interviewer**: Uh-huh.
> **Mito**: then I think 'Well it's about time now to look this up because it's worth it'
> **Interviewer**: Ok.
> **Mito**: I mean, I actually look up quite a lot, because I always get ah, see, new words again, where I think to myself 'Ah, I've actually already asked myself several times, x that I don't know that', yes.

The last feature which characterizes words that participants reflect about or use VLS for is that they spark their interest. For DJ (SF01: 343-354) that means that unknown words get stuck in his head and really annoy him until he finds out what they mean. This sentiment of being annoyed by unknown words is also well known to Walküre (SF01: 451). Lia (SA02: 398) reports sometimes wanting to know what exactly a word means, but all three of them do or cannot explain why that is the case. Vanessa's attention, on the other, is attracted by words that sound or look interesting:

> **Vanessa:** Ich achte nicht so sehr darauf, dass ich jetzt alle Wörter versteh,
> **Interviewer**: Vollkommen ok.
> **Vanessa**: aber wenn jetzt ein Wort so interessant klingt oder ausschaut, dann schau ich nach, ja. (SD01: 451-453)

> **Vanessa**: I do not pay a lot of attention to understanding all the words
> **Interviewer**: Completely fine.
> **Vanessa**: but well if a word sounds or looks interesting, then I look it up, yes.

While it does not become clear at all from the focus group interviews what it is that attracts participants' curiosity, these four examples suggest that some words hold an inherent interest although what captures their attention is likely different for different students.

On the whole, the quantitative findings on strategies to discover the meaning of unknown words encountered in EE input are supported by participants in the focus groups. Guessing from context is again positioned as the default option and additional VLS such as thinking about other languages and using online dictionaries are only used if it fails. Analysing word parts and identifying part of speech are seen as too difficult and time-consuming to be used regularly and asking others for help is often not feasible for reasons of proximity and language competence. In addition, the interview data show that in order to make the use of further resources or strategies in addition to contextual inferencing worthwhile, unknown lexical items need to have

a least one of the following three characteristics: they need to be crucial for understanding the content *and* difficult to infer from context, they have been encountered repeatedly during EE activities, or they somehow arouse learners' interest.

7.4.2 Remembering new words

The second aspect related to students' accounts of lexical learning regards the extent to which they remember words acquired from EE activities and what helps to remember them, the latter being clearly related to participants' beliefs on what helps or hinders learning from EE (see section 7.3.3).

Maria and Jane (SC01: 533-537) emphasize that memorizing new words, from EE activities and in general, is a difficult endeavour. This sentiment is shared by their classmate Paul, who argues that learning is complicated by the fact that most unknown words are not encountered very often (see also section 7.3.3):

> **Paul**: Aber die meisten Wörter, die man jetzt nicht so kennt, die begegnen einem auch nicht so oft. Ich mein, wenn sie dir bis jetzt nicht (wirklich begegnet sind), werden sie auch vielleicht nicht so oft begegnen, deswegen [...] (SC01: 540)
>
> **Paul**: But most of the words that you do not know, they also do not come up that often. I mean, if you (haven't really encountered) them until now, perhaps you won't encounter them that often, that's why [...]

While this statement is very insightful and probably true for a large proportion of low-frequency vocabulary, other participants are more optimistic and maintain that they do remember new words picked up from EE activities.

Jane (SC01: 376) argues that hearing and seeing words several times helps her to remember them and such a repetition effect is also highlighted by Anna and Mito (SD01: 223-323), Karl (SE01: 475-478), Walküre (SF01: 479-482) and Louise (SG02: 536). Another argument is that looking words up in dictionaries helps to deal with different word meanings, and thus understanding them in different contexts (Maria, Jane & Paul 547-550), and to know the exact meaning, which makes memorizing them easier (Susi SD01: 218-222, Johannes SG02: 538). Kira (SA02: 414) also thinks that meeting a word in different contexts helps her to remember it, whereas for Kirito (SE01: 473) or DJ and Walküre (SF01: 469-477) it depends on how important a word is for a scene or the plot. In addition, immediately trying to actively use new lexical items is a memory strategy advocated by John M. and Jane (SC01: 701-707).

Some students appear to believe that inferring the meaning of a word from context helps to remember it (Marie SG02: 530, Anna & Mito SD01: 223-323), while others (Elisa, Louise & Johannes SG02: 533-535) are not so sure. These differences in opinion relate to different notions of what it means to know a word (see also section 7.3.1) and to the question what the goal of memorizing a new word is. Elisa (SG02: 550-554), for instance, argues that if she learns a word by inferring it from context, she cannot answer questions about its exact meaning, even though she can use it productively. While the latter part of her claim might be considered questionable, a similar thought is expressed by KingKong:

> **KingKong**: Bei mir ist es irgendwie so. Ich kann mir die Sachen halt ziemlich lange merken und weiß halt auch immer was sie heißen, aber (1) ich kann's halt nicht so gut erklären dann, wenn ich's halt nur aus dem Zusammenhang kenne. Also wenn ich ein Wort kennenlern und dann sagt mir jemand ‚Erklär's mir', dann kann ich das halt nicht. (SF02: 486)
>
> **KingKong**: For me it's somehow like that. I can remember things for a pretty long time and I also always know that they mean, but (1) I can't explain it well then, if I only know it from context. So if I get to know a new word and then someone says 'Explain it to me', then I simply can't do it.

KingKong agrees with Elisa that learning words from context is not sufficient for being able to explain their meaning, but he asserts that he understands them. Although he talks about "a word" in the quote above, multiword units and idiomaticity could also play a role here because idiomatic meanings are often even more difficult to explain. Similar thoughts on learning from context are expressed by Jane (SC01 551-553) with regard to translation and Anna and Mito (SD01: 223-323) also concur that inferring from context leads to learning, but it may not be enough for precise translations. Consequently, participants' evaluations of whether inferring from context is useful seem to depend on their conception of the learning aim: if it is enough to roughly remember what a word means, then it is seen useful, but if the aim is being able to explain or translate a given word's meaning, it is regarded as insufficient.

In addition to describing their experiences with remembering new words encountered in EE input, students were asked whether they could think of any concrete examples of learning vocabulary from EE. In the following, all instances of vocabulary learning described in the focus groups are provided together with information on the source of learning, if participants could remember,

Several participants provide examples of new English words acquired from series, films and video clips. John M. (SD01: 668-667) learned the word *turmoil* from a video about the Great Depression in the USA and Paul (SD01: 684-694) picked up *vanguard* from the fantasy series *Game of Thrones*. During the interview Paul checks the history of his dictionary app and reports that he also recently looked up *binge drinking* and *appendix* (referring to the body part) but cannot remember where he came across these words. Jane (SD01: 678 – 683) remembers finding out about *concussion* in the context of a YouTube video of Americans who engage in funny competitions and Elisa (SG02: 613-616) looked up the word *inception* because she wanted to understand the title of the movie *Inception*. However, participants do not only remember picking up the meaning of new words, Kira (SA02: 318-323), for instance, learned the pronunciation of *hovercraft* from the movie *The Hunger Games*.

Other students report learning words from written text: Anna (SD01: 292-301) read a book about economy where she encountered the word *fiscal* and Kirito (SE01: 495-499) learned *perennial* from a novel. Elisa (SG02: 603-609) frequently saw the word *procrastination* in posts on Instagram and then looked up its meaning. Concerning songs, only DJ (SF01: 564-577) provides an example: he remembers finding out that *whips* can also be used as a synonym for *cars* from one of his favourite rap lines, which reads "I got more whips than a runaway slave" and how he then realized the different ways in which this line can be understood.

Students also learn words from peers as the next two examples show. Louise (SG02: 589-601) helped a friend create a character for the fantasy role-playing game *Dungeons & Dragons* and acquired the word *deception* in the process of doing so. Pinguin (SE01: 479-493) reports learning the word *discombobulated* from his friend Karl's cousin, which he then taught his other friends because he thinks that it is really funny.

In addition to these lexical items, a few students mention picking up swearwords and youth slang such as the abbreviation *SMH* for *shaking my head,* mostly in the context of social media. However, the examples described above clearly suggest that teenagers do not only learn informal language from their EE activities, but that they also encounter and memorize relatively formal and infrequent words, such as *turmoil* or *perennial,* or specialized vocabulary, such as *fiscal* in the context of economy. This finding again highlights that specialized interests result in highly individualized EE environments which can lead to different learning gains, as mentioned in context of niche activities in section 7.2.1.

In sum, the focus groups allow for valuable insights into students' perceptions of several aspects of lexical learning. Participants' accounts of strategy use position guessing from context as the default option which, if considered necessary, is supplemented by thinking about other languages and using online dictionaries. Other VLS like analysing word parts or identifying the part of speech are regarded as requiring too much time and effort and the social strategy of asking others is generally seen as infeasible because a competent other person is frequently not at hand. These findings are in line with the results of the quantitative strand, which are also evaluated positively by participants. Concerning the characteristics of words for which participants think the effort of using VLS other than guessing from context worthwhile, three features emerge as significant: students are likely to look up unknown words that are (1) crucial for understanding the content and difficult to infer from context, (2) encountered repeatedly during their EE activities, or (3) interesting in some way and thus capture their attention. Participants argue that after having discovered the meaning of new English words, they do remember them, but opinions on what helps to remember novel lexical items vary: some stress that repeated encounters or the salience of the word in the context of the encounter are crucial, while others think that finding out the exact meaning from a dictionary or actively using new words helps to memorize them. The question whether words inferred from context are remembered well results in very different evaluations, which appear to be related to participants' conceptualizations of the aim of learning and of what it means to know a word. Finally, examples for lexical learning from EE activities provided in the focus groups highlight the fact that students come across a wide range of vocabulary outside school and can acquire rather formal and infrequent words through their leisure time activities.

7.5   The relationship between in- and out-of-school English

Focusing on the potential interaction between English practices inside and outside the classroom, this final section endeavours to answer the following research question: *How do participants describe the relationship between their out-of-school English activities and their English lessons at school?* (RQ 5g). Participants' perceptions of possible links between in- and out-of-school English are described first before addressing the question whether knowledge acquired from EE helps or creates problems in English lessons.[206]

Figure 7.5: Word cloud of the 100 most frequent types in the main category THE RELATIONSHIP BETWEEN IN- AND OUT-OF-SCHOOL ENGLISH

Three different views on the relationship between in- and out-of-school English were identified in the focus group data. The majority of nine participants detect no link between their lessons at school and their leisure activities, while four participants see a connection in form of mutual influence.[207] Furthermore, six students explicitly state that in their opinion the basic English proficiency acquired at school forms the foundation for engagement with EE. It is worth noting that this last belief does not stand in contrast to the second opinion; in fact, three of the four participants who think that their in- and out-of-school English practices influence each other also mention that their lessons at lower secondary level equipped them with the basic language skills needed for using English during their spare time.

As mentioned above, nearly a third of all students participating in the focus groups see their English teaching at school and their own EE practices as completely separate entities, like Lukas in the quote below:

> **Lukas**: Nein, also ich wollt nur sagen mich beeinflusst's eigentlich überhaupt nicht. Also (1) ist bei mir eigentlich komplett getrennt.
> **Interviewer**: Ok. (2)
> **Lukas**: Ich verwend's halt einfach nicht so viel in der Schule was ich beim was ich sonst außerhalb vom Unterricht auf Englisch machen würde. (SG02: 634-636)

---

[206] As mentioned in section 5.4.6, data on participants' views on whether in- and out-of-school English should be integrated were also collected in the focus groups; however, for reasons of scope it was finally decided not to include an additional pedagogical focus in this thesis; therefore, these data will be presented at a later date.
[207] 17 of 30 participants in the focus groups explicitly expressed their view on this topic.

> **Lukas**: No, well I just wanted to say that actually it doesn't influence me at all. So (1) for me it actually is completely separate.
> **Interviewer**: Ok. (2)
> **Lukas**: I simply don't use a lot at school of what I what I would otherwise do in English outside of lessons.

KingKong (SF01: 265-271) agrees that what they do with English at school is totally different from what he does outside school and Anna (SD01: 190-193) thinks that what her class learns at school is completely irrelevant for her free time. Emma adds another dimension in the statement reported below:

> **Emma**: Für mich hat's überhaupt keinen Zusammenhang.
> **Interviewer**: Ok.
> **Emma**: Also wir machen ganz was anderes (2) im Englischunterricht als ich draußen mache, also (2) ich übe draußen Englisch draußen,
> **Interviewer**: Draußen,
> **Emma**: in der freien Welt. @ (SD01: 485-492)

> **Emma**: For me there is no connection at all.
> **Interviewer**: Ok.
> **Emma**: Well we do something totally different (2) in our English lessons than I do outside, well (2) I practise outside English outside,
> **Interviewer**: outside
> **Emma**: in the free world. @

Many of Emma's peers agree that they learn more outside school nowadays, as we shall see below; however, the different foci of English practices in and outside school are not necessarily evaluated negatively.

In contrast, four participants are convinced that in- and out-of-school English affect each other; Mito even thinks that this is perfectly obvious, as he explains in the following extract:

> **Mito**: Ja, ich mein, ich glaub nicht, dass man jetzt sagt in der Freizeit ‚Aha, ich hab jetzt ein Wort, das haben wir zum Beispiel in der Schule gelernt, aber da hab ich einen Trennstrich und ich schau's jetzt nochmal nach', also was man in der Schule lernt, also das das benutzt man vielleicht zu Hause und wenn jetzt man wenn man zu Hause ein Wort kennengelernt hat, was nie in der Schule gefragt wird, dann dann also ich glaub schon, dass, weil man hat ja einen Englischwortschatz und jetzt nicht den von der Schule und von zu Hause (SD01: 509)

> **Mito**: Yes, I mean, I don't think that someone says in their free time 'Ah, now I have a word here that, for instance, we learned at school but I have a separating line there and I look it up again', so what you learn in school, well that you use that at home perhaps and if someone has gotten to know a word at home, that isn't ever asked for in school, then then well I do think that, because you have just one English vocabulary and not one for school and for at home.

In a similar vein, Louise (SG02: 643-645) argues that using structures acquired at school or encountering them again in out-of-school contexts helps to consolidate her knowledge, for instance, in the case of vocabulary. Kira (SA02: 409-414) describes the example of coming across a word during her spare time, for example in a song, and then seeing or hearing it again in her English lessons and thus also concludes that there definitely is mutual influence. Kirito presents his point of view in a very poetic manner:

> **Kirito**: Also meiner Meinung nach, also in der Schule, auch wenn man's nicht will, wird man gezwungen, die Tür zu der englischen Welt aufzuschlagen, die englische Welt wird vor allem aufgeschlagen, die Tür dazu,
> **Interviewer**: Ok.

> **Kirito**: und man selbst wandert durch und nimmt Wissen von dort mit und bringt's dann wieder in die Klasse zurück. (SE01: 512-514)
>
> **Kirito**: Well in my opinion, so at school, even if you don't want it, you are forced to open the door to the English world, the English world is being opened up, the door leading to it,
> **Interviewer**: Ok.
> **Kirito**: and you wander through and take knowledge from there with you and bring it back to the classroom again.

Kirito does not only summarize very nicely the mutual influence between in- and out-of-school contexts, which both contribute to language development, he also suggests that this relationship is initiated at school. This aspect is at the heart of the third perception, which sees school as the basis for further language learning from engagement with EE.

Six students explicitly express the view that a basic level of English language proficiency is acquired at school, which is then expanded through EE activities; however, more participants seem to agree with it in the interviews (e.g. John M. & Jane SC01: 730-748, unidentifiable students in SD01: 494-497 and SA02: 440). Kira is a typical representative of this opinion, explicitly voicing it in the quote below:

> **Kira**: Also ich find das Englisch, das Schulenglisch, hat mir zum Beispiel den Start gegeben um dann außerhalb der Schule Englisch gut zu verstehen. Also die Grammatik hätt ich nie außerhalb der Schule gelernt.
> **SX-f**: xx Unterstufe.
> **Kira**: Und das war quasi die Basis und jetzt außerhalb der Schule baut man's halt aus. (SA02: 439-441)
>
> **Kira**: Well I think that English, the school English, has for example given me the start to then understand English outside of school. I mean, the grammar I would never have learned outside school.
> **SX-f**: xx lower secondary.
> **Kira**: And that was the basis so to speak and now outside school you build on it.

In addition to positioning English teaching at school as the foundation of extramural English, Kira introduces a point mentioned by several students, namely that grammar in particular was learned and needs to be learned at school. Other students like Vanessa (SD01: 494-496) and Paul SC01: (729-731) also see teaching at school as having laid the groundwork for out-of-school engagement with English. Similar to Kirito above, Susi (SD01: 499-501) emphasizes that at school everybody is forced to learn English and those who are interested can then do more with it, but everybody has to learn a minimum, which she appears to evaluate positively. Her classmate Mito (SD01: 511-514) agrees that first one needs to understand the basic principles before one can use English independently, but he adds that now at upper secondary level they do not learn a lot of new things at school, especially in terms of grammar. This view that few new structures are learned at school after lower secondary level is shared by Jane and Maria (SD01: 736-740) as well as Lia (SA02: 433-435), who criticizes that all they do now is talk about various topics and learn *"Gebildetenenglisch"*, a kind of more sophisticated English used by educated people, which she clearly views negatively.

The perception that at upper secondary level not much new is learned at school goes hand in hand with a belief that once a basic knowledge of English has been acquired through lessons, more is learned outside the classroom. This opinion is given by Lia (SA02: 422), but it is also expressed by several other participants including those who see no link between English inside

and outside of school. In group SD01 Mito (SD01: 185) argues that while school is helpful, it is only responsible for a small part of their English proficiency, while the larger part is acquired during their spare time, to which his classmates Anna and Emma (SD01: 480-484) heartily agree. Later, Anna introduces another aspect to this group's discussion by arguing that engagement with EE is even necessary to keep up with one's peers in terms of English proficiency:

> **Anna**: Und ich mein, wenn man schon in der ersten und zweiten mit diesen <ENGLISH> he she it <ENGLISH> (noch) nicht klarkommt, dann, ich weiß nicht, man ist, es ist einfach so, man muss sich einfach wirklich teilweise dafür auch interessieren, wenn man das Mindeste macht, gut dann kommt man durch, aber so wenn man auch in der Ersten, Zweiten, Dritten nur das Mindeste gemacht hat, dann kommt man in der Sechsten einfach nicht mehr weiter,
> **SX-f**: Ja.
> **Anna**: dann kann man einfach gar nichts mehr in der Sechsten.
> **Interviewer**: Mhm.
> **Anna**: Und ich weiß nicht, das ist auch in unserer Klasse so, das sieht man auch bei vielen Leuten bei uns. (2) Und da muss man einfach irgend- irgendwie muss man in der Freizeit einfach irgendwas damit zu tun haben (SD01: 503-507)

> **Anna**: And I mean, if you cannot (yet) cope with <ENGLISH> he she it <ENGLISH> in first and second form, then, I don't know, it is just like that, you simply have to be interested in it partly, if you do the minimum, well, then you pass, but if you only did the minimum in the first, second, third form, then you just don't progress in sixth form.
> **SX-f**: Yes.
> **Anna**: and then you just can't handle anything anymore in sixth form.
> **Interviewer**: Mhm.
> **Anna**: And I don't know, it's also like that in our class, you just see that also with a lot of people among us. (2) And then you simply have to have some- something to do with it in your free time

Anna's statement may be an example of an individual belief not held by others, but it is noticeable that for students in group SD01, as well as in all other focus groups, the existence of a link between out-of-school engagement and level of English proficiency is self-evident. In group SG02 (244-250), Marie even argues that she acquired certain language structures outside school before they were the subject of lessons, thus partly contrasting others' opinion that school lessons provided the basis for their English proficiency. To sum up, there is a majority of students who do not see a link between the English practices inside and outside school, but there also is a strong belief that engagement with EE is beneficial for language development and that once a basic English proficiency has been acquired at school, more can be learned from using English outside school.

Furthermore, in relation to their current English lessons EE can either be seen as helpful or as a problematic influence. Some students argue that the positive effects of EE on language learning are useful for the school subject, like Anna (SD01: 503-507) above. For instance, Keanu (SA02: 290-292) uses EE for test preparation and Elisa (SG02: 625-629) argues that she has better comprehension skills and a feeling for what sounds right in English. DJ and Walküre (SF01: 590-611) add that it helps them with pronunciation and text writing, especially if they can use vocabulary they have learned outside school. At the same time, students also identify problematic aspects with regard to using knowledge acquired outside school in their English lessons. Two examples of learning wrong grammatical structures from songs provided by

KingKong (SF01: 612-623) and Marie (SG02: 370-372) have already been discussed in section 7.3.3. The realization that he picked up a wrong structure leads KingKong to the following conclusion:

> **KingKong**: und seitdem versuch ich das immer zu trennen, dass ich halt nicht irgendwelche solche <ENGLISH> slang </ENGLISH> -Fehler einbau. (SF01: 623)
>
> **KingKong**: And since then I try to always keep it separate, so that I don't make any such <ENGLISH> slang </ENGLISH> mistakes.

Similarly, Marie and Elisa (SG02: 296-309, 313-328) also problematize the influence of EE on their language use in school, although both of them argue that they learn a lot of English outside school (see above). They mention that they have to be very careful not to use slang expressions or swearwords when speaking in class and that their style of writing is often rather colloquial, which sometimes is an issue depending on the type of writing task they have been set.

To conclude, almost a third of all students in the focus groups see no relation between their lessons at school and their English practices in other contexts, although they do not necessarily evaluate this situation negatively. In contrast, only four participants see a connection in form of mutual influence. Another common perception is that English lessons provide the basic proficiency required for using English independently; therefore, school is seen as the foundation for engagement with EE. However, participants argue that once this basic level of proficiency is reached, more is learned outside school because in their view not much new is learned in English lessons at upper secondary level, particularly in terms of grammar. Using knowledge acquired through EE activities in English lessons is regarded as helpful by most participants, but sometimes it can also create problems as suggested by individual examples.

7.6   Summary

This chapter presented the findings of the qualitative strand based on six focus group interviews with 30 participants. Analysis of the interview data has shown them to be an invaluable source of information, revealing diverse, multi-layered and sometimes contradictory beliefs, as is typical of lay theories. Results show that teenagers hold strong views on a number of issues relating to informal (and formal) language learning and indicate that engaging in discussions with learners can lead to unexpected insights and prove enormously fruitful for research purposes.

Findings reveal that knowing English is evaluated as important across all groups, but assessments of the role of English in the lives of young Austrian vary: some see knowledge of English as absolutely essential for teenagers such as themselves, whereas others argue that within Austria it is an additional benefit rather than a necessity. In statements comparing English to other (foreign) languages it is, however, consistently described as the most important language next to the participants' first language(s), which emphasizes its special status. The main reasons given for the significance of English are its role as 'the universal language', as one participant characterized it, and its use in leisure time activities. The first reason relates to the current position of English as a the most widely used lingua franca, which has perceived

implications for international connection and participants' future, whereas the second clearly relates to extramural English.

Regarding engagement with extramural English participants corroborate the quantitative results on the most frequent EE activities: the fact that music is one of the most popular activities is unquestioned since almost all songs teenagers listen to are in English and series, films and videos are presented as very common points of contact with English in all focus groups. In addition, participants read English quite a lot, mainly in online environments, but book and other print texts are also discussed. In contrast, little writing is done and speaking, which is not discussed much, appears to be limited to communication with non-German-speaking friends and relatives. Four boys present an exception because they appear to play (multiplayer) games quite frequently and therefore make use of both written and spoken English in different forms of in-game communication. In addition to these more general results on engagement with EE, the cases of four participants who engage in highly specialized niche activities exemplify the range of activities for which Viennese adolescents use English.

Discussions of time spent with EE in the focus groups are of particular interest because the estimate established on the basis of the EEOLD data appeared incredibly large at first glance. Similarly, many participants first expressed disbelief in the interviews, but upon closer reflection on their own EE practices most of them came to the conclusion that a mean EE time of approximately four hours a day is plausible. In addition, a third of the participants immediately confirmed four hours as a reasonable amount of time with some presenting a convincing case that for them the estimate is even too low. Two factors that emerge as important contributors to the amount of time spent with EE are the role of music and the possibility of engaging with EE 'on the move' through the use of smartphones. In addition, participants point out that EE activities do not always involve a conscious focus on language and that some can also be done simultaneously to other tasks, which is particularly true for listening to music.

Concerning the reasons for engagement with EE, the qualitative data indicate that a preference for original versions, the aesthetic qualities ascribed to the English language, and the availability of content are the most important grounds for the use of English in leisure activities. While aesthetic reasons appear to be more important overall, availability plays a role in relation to the exclusive availability of certain (online) contents in English, the delay in the availability of dubbed or translated versions, and the possibility of accessing a wider pool of information through the use of English. A further reason given by several participants is the aforementioned use of English for communication with non-German-speaking friends or relatives.

The second major focus of the focus group interviews are participants' views on learning from EE in general and on vocabulary learning in particular. The results show that a large majority of participants believe that they benefit from their EE activities in terms of language development. The students identify a wide range of aspects that can be learned from EE: vocabulary and other lexical features are named most frequently followed by pronunciation and procedural knowledge of English which leads to more natural language use. Beliefs on what

helps or hinders learning from EE are diverse and strongly related to participants' personal experiences, but many appear to be quite strongly influenced by a conviction that English is best learned through communication with or exposure to L1 speakers. Conversely, the most frequently mentioned problem is the possibility of learning 'wrong' language from other L2 speakers; in addition, incorrect grammatical structures or slang terms used, for instance, in songs are also seen as problematic. Overall, however, participants name fewer problems than benefits of EE for language learning and regard a positive effect on motivation, repeated encounters and a familiarization effect as contributing to their language development. Concerning the learning potential of individual EE activities, participants' evaluations differed widely. On the whole, productive language use through speaking and chatting as well as reading are seen as activities with high learning potential followed by audiovisual media. The learning potential of games and music was the subject of heated discussions, most students see limited possibilities for learning from songs, whereas the potential of gaming depends on the type of game played in their view.

In line with the aims of the study, the acquisition of new words from EE input was discussed in more detail with most discussions focusing on word meaning. Participants' accounts indicate that guessing from context is the default strategy used to discover the meaning of unknown lexical items and, if needed, they use comparisons to other languages or online dictionaries as additional strategies. However, only certain new words are worth the effort of using strategies in addition to contextual guessing: in order to receive further strategic attention, unknown lexical items either need to be crucial for understanding the content *and* difficult to infer from context, or encountered repeatedly in EE input, or they need to attract participants' attention in some way, mostly likely through formal properties. Participants' views on what helps to remember the meaning of new words once their meaning has been discovered vary more widely from repeated encounters to productive use. Yet, some vocabulary is clearly remembered because participants are able to provide a number of examples of lexical words and phrases learned from EE activities.

The final focus of the focus group interviews concerned the relationship between English inside and outside the school context. Most participants see no connection between these spheres, while some point out that the two contexts naturally influence each other and others again see a different link because they argue that English lessons at school provide the foundation for extramural English activities. At the same time, many participants argue that by now their out-of-school engagement with English has contributed more to their knowledge of English than teaching at school. This view appears to be related to a perception that after basic vocabulary and grammar have been acquired at lower secondary level, English lessons at upper secondary are repetitive and largely irrelevant. Hence, for the teenagers in the focus groups English teaching at school often carries negative connotations, while using and developing their English proficiency in 'real-world' activities in extramural contexts is evaluated positively.

8 Discussion

Adding to the existing body of research on extramural English and its relation to vocabulary learning by exploring the informal practices, lexical knowledge and language learning beliefs of upper secondary school students in Vienna, Austria, this thesis contributes insights of a conceptual and methodological nature as well as new empirical findings. This chapter thus integrates the results of the quantitative and qualitative strand in the tradition of mixed methods research (see section 5.2.1), but at the same time discusses the wider implications of the study for the field. First, conceptual and methodological conclusions are presented in relation to extramural English, the use of mixed methods, and vocabulary measurement (section 8.1). In a second step, empirical findings bringing together the results of the quantitative and qualitative strand of the MMR project are presented and discussed in light of previous research (section 8.2).

8.1 Discussion of conceptual and methodological insights

The conceptual and methodological contributions of this project fall into the three broad areas of extramural English and related conceptualizations of informal language learning and use, the adoption of a fully integrated MMR design, and issues in relation to L2 vocabulary measurement. Below, each of these areas is discussed in turn beginning with EE as the central concept of this thesis.

As set out in Chapter 2, this study forms part of an emerging research field interested in language learning through informal leisure activities outside educational institutions. Owing to different research backgrounds and foci, the research community currently uses a wealth of different terms for this object of study (e.g. Benson & Reinders 2011b). Hence, the first step towards a clear conceptualization of the area of investigation in the present study included an attempt to disentangle and define some of the more prominent conceptualizations currently used in the field: *extramural English* (EE, Sundqvist 2009a)*, language learning beyond the classroom* (LBC, Benson & Reinders 2011b)*, the online informal learning of English* (OILE, Sockett 2013, 2014) and *informal digital learning of English* (IDLE, Lee 2019a, 2019b; Lee & Dressman 2018). A close analysis and comparison of these four conceptualizations (see sections 2.1 and 2.2) has highlighted important differences in terms of their definitional focus, original motivation, perspective on learning, and scope.

First, the four concepts differ with regard to the focus of their definitions: while EE and LBC use spatial definitions, OILE and IDLE are defined in relation to media. Second, in seeking to establish a framework for the emerging field, LBC clearly has a theoretical motivation, whereas EE, OILE and IDLE originally represent more empirical perspectives. Third, and perhaps most crucial, these approaches also differ in relation to their perspective on learning: while LBC, OILE and IDLE are specifically interested in and explicitly refer to language learning, EE does not posit learning as a given; rather, it presents a more inclusive perspective that investigates practices of language use which may or may not lead to language learning. In addition, both EE and LBC

do not restrict their focus to one type of learning, whereas OILE and IDLE specifically focus on informal learning. Clearly, the vast majority of language use and learning not related to formal educational institutions is of an informal nature; however, EE and LBC recognize the fact that in some cases language learning outside classrooms and schools may be intentional on the part of the learner. At the same time, the two conceptualizations clearly differ in terms of their inclusiveness in this respect: while Benson's (2011) model explicitly includes extracurricular settings and instructed, teacher-directed learning as part of LBC, Sundqvist and Sylvén (2016) specify that EE occurs outside formal educational settings and emphasize the learner-drivenness of EE activities.

The differences highlighted so far thus clearly show that the EE, LBC, OILE and IDLE are very different in scope: while OILE and IDLE present narrower foci on informal learning through specific media and can be differentiated solely by the (lack of) inclusion of offline contexts, EE and LBC are much wider in scope. As mentioned above, LBC represents the broadest approach in terms of settings and forms of learning included in the conceptualization; however, its primary interest is language learning. Hence, from a different perspective EE can be considered as having a wider scope than LBC because it is interested in all language practices outside formal education, even though they do not necessarily always entail language learning.

This comparison, which highlights different foci and different strengths of the four conceptualizations, indicates that for a project aiming to gain as comprehensive an overview as possible of a new research environment and participant sample, EE is the more obvious choice. In addition, the exclusive focus of LBC on language learning, which is foregrounded in Benson's (2011: 13) definition of modes of practices as "routine pedagogical processes", is problematic from an empirical perspective. As discussed in section 2.1, it is very difficult if not outright impossible for researchers to distinguish between routine pedagogical processes and other social practices because such a differentiation would presuppose a priori knowledge of learners' intentions when engaging with the L2 in informal, out-of-school settings. For this reason, I have proposed a new working definition of modes of practice, which builds on and extends Benson's (2011) original definition:

> A mode of practice is a set of routine social practices which are located in and deploy features of a particular setting, involve target language use and have pedagogical potential.

This more inclusive definition can be operationalized more easily for research purposes in real-world contexts and is more compatible with EE as the primary concept used in this project. Building on the work of Sundqvist (2009a) and Sundqvist and Sylvén (2016), I have thus conceptualized the object of investigation for the present study as extramural English among 15-to 16-year-old academic secondary school students in Vienna, defined as English-language activities which take place outside the walls of educational institutions during learners' leisure time and which are learner-driven and typically voluntary and informal (see section 2.4).

The empirical findings of the MMR study support this conceptualization and the use of EE as the central concept because results firstly show that English-language practices of Viennese teenagers are not confined to online environments and digital media, and secondly that language learning is not the primary purpose for engaging with English outside school. Although the most frequent EE activities make use of digital media and are carried out in online contexts (see section 6.3.1), a narrower conceptualization such as IDLE or OILE would have excluded less frequent niche activities such as reading books or print media, writing diaries, creating comics and manga, or participating in English-speaking sports teams (see sections 6.3.1 and 7.2.1). This study thus suggests that adopting a broader perspective including more traditional media and non-media-related activities in addition to digital media activities is worthwhile to capture a full picture of the immense variety of L2 English learners' out-of-school exposure. In addition, both quantitative and qualitative data indicate that adolescents have a variety of reasons for using English in their spare time among which improving their language skills is only one of several factors and not the dominant one (see sections 6.3.3 and 7.2.3). This outcome that learning is not the primary purpose of informal English activities and the resulting difficulty of differentiating between informal extramural language use and language learning lends further support to the more inclusive EE perspective and as well as to the new working definition of modes of practice proposed in section 2.1.

In addition to clearly delineating and differentiating extramural English from other related concepts in the field, Chapter 2 also presented an extensive overview of previous research. The narrative review along the five meta-analytic dimensions of time, location and context, researchers' background and research interests, theoretical concepts and content focus, and research design (section 2.3) shows that the emerging research area has received increased attention over the last decade so that by now informal L2 learning outside educational settings has been investigated in highly diverse environments across all continents. The vast majority of studies have however focused on EFL settings in Europe and Asia, while other contexts and especially languages other than English have not yet received sufficient attention. Within the European context, most studies have been carried out in subtitling countries in which children are exposed to English from an early age onwards, although research on non-subtitling countries such as Austria is gradually increasing.

The review further indicates that the backgrounds and interests of researchers working in this field are highly diverse, which is reflected in the diverse terms and conceptualizations used (see Table 2.2). In some studies the object of investigation is not clearly defined and remains rather vague; hence, the field would certainly benefit from theoretical consolidation and further development of models such as Benson's (2011) framework, which, although preliminary, has proved very influential. Fewer terms and clearer conceptualizations would help reduce terminological confusion and increase comparability between studies to move the field forward as a whole. In this respect, looking to neighbouring discourses on informal learning in relation to science and other areas as well as digital learning (see Chapter 2) may be a worthwhile

endeavour to see whether frameworks and conceptualizations used in these fields could fruitfully be transferred to research on language learning.

Finally, the review also suggests that further empirical research is needed in several areas: first, studies using rigorous research designs are needed to provide further evidence on the effects of extramural activities on language development; second, more in-depth investigations are necessary to understand how learners engage with language in such informal activities, and third, carefully planned qualitative studies are needed to better comprehend learners' perspectives and to understand whether, to what extent and how out-of-school language learning can be connected to in-school learning and teaching (see also Reinders & Benson 2017; Schmitt 2019; Sockett 2014).

At the same time, it has to be acknowledged that studying a real-world phenomenon such as extramural English under naturalistic conditions presents several difficulties. As pointed out by Sockett (2014: 110) about the related concept of OILE "the private and individualised nature of OILE activities" makes data collection more challenging because OILE or EE activities can rarely be observed directly. A further complication is the impossibility to include a control group because the researcher cannot exert any control over engagement with EE during participants' leisure time, which renders the use of a classic research design including experimental and control groups infeasible. In addition, leisure trends and possibilities of access to EE change relatively quickly; thus, results are necessarily limited to the specific time and context of data collection and need to be interpreted with these in mind (Lai 2015; Lai, Zhu & Gong 2015). While this is true of many applied linguistic or SLA studies, it is perhaps more important to keep this fact in mind in studies on informal language learning due to the volatile nature of the current media landscapes.

Despite such difficulties, the present study endeavoured to gain a detailed picture of extramural English among Viennese teenagers from several perspectives. The conceptualization of EE as presented in Chapter 2 was operationalized in a carefully designed and extensively piloted questionnaire, the EEQ, to collect information on the type of contact with English in out-of-school contexts as well as the frequency of engagement (see section 5.3.3.1). A detailed multi-item scale elicits information not only on the frequency of possible EE activities, but also on the channels and devices used to engage in these activities. These data on access to EE allow linking the results of the present study to media surveys such as those presented in Chapter 4 and suggest that adolescents' preferred access points and media activities are strikingly similar across different languages (see sections 6.1.3 and 6.3.1). Based on this finding that media use in L1(s) and English as an L2 are essentially the same, at least in the Austrian context investigated, future studies could make use of less detailed EE taxonomies and focus on other aspects. In addition, the EEQ collects data on several additional variables to allow the analysis of sociological and linguistic influencing factors on EE. Variables such as socioeconomic status, the number of home languages, or demographic data were operationalized drawing on the expertise of large-scale

survey studies (e.g. OECD 2014) to allow comparisons with these. In addition, further variables such as attitudes towards or awareness of English were also included for exploratory purposes.

Following the example of previous studies (Olsson 2012; Olsson & Sylvén 2015; Sundqvist 2009a), the questionnaire data are complemented by a structured online language diary, the EEOLD, which provides information on time spent with EE and thus on the amount of contact (see section 5.3.3.2). The EEOLD emerged as the most problematic instrument used in the present study because it was administered online rather than in the presence of the researcher and the requirement of anonymity precluded personalized reminders. Although problems with diary instruments have also occurred in other studies (e.g. Olsson 2012) and the MMR design used in this study helped to remedy some issues with regard to the language diary data as will be discussed below, further methodological development is certainly needed in this area. To improve the response rate, future research could either directly involve teachers, which was avoided in the present study so as not to contribute to teachers' already immense workload, or use more advanced technological options such as apps that could remind participants via notifications without the researcher needing access to personal contact information.

Finally, another perspective on EE was offered by students' qualitative accounts of their experiences with using English for informal leisure activities in the focus groups. The qualitative data allow insights into teenagers' attitudes and beliefs with regard to the English language and their use of English as well as into their lay theories about learning English both inside and outside school, which have proved to be an invaluable source of information. Thus, the quantitative data collected in this study offer information on the type, frequency and amount of EE as well as access points and potential influencing factors, while the qualitative data provide complementary information on learners' experiences with and beliefs about EE and its learning potential. Hence, even if concentrating solely on EE, the benefits of mixed methods research become immediately obvious, but the use of a fully integrated MMR design had several further advantages.

As envisaged from the beginning and set out in Chapter 5, mixing methods led to benefits in terms of confirming, complementing and explaining the results of one research strand through the other. In several cases, for instance concerning the reasons for using English or the use of vocabulary learning strategies, quantitative and qualitative instruments produced similar outcomes, which confirmed each other. Moreover, the quantitative and qualitative perspectives clearly complement and enhance each other; for example, in addition to confirming the quantitative findings the qualitative data also reveal why certain VLS are used more extensively than others. However, the most important effect of using MMR in the present study was the confirmation and explanation of seemingly implausible results. As mentioned above, data collection with the EEOLD was subject to limitations and therefore the results concerning time spent with EE could have been called into question, particularly since the amount of time appeared to be inconceivably large at first (see section 6.3.2). The possibility of discussing findings such as these in the follow-up focus group interviews emerged as an invaluable

opportunity for member checking and actually showed that the estimates established using the EEOLD are credible. Consequently, the careful planning and rigorous implementation of mixed methods at different stages of the research design clearly enhanced the trustworthiness of data through triangulation and helped to arrive at a richer, more detailed analysis that includes different angles on the area of investigation.

In addition to these overall benefits, the study also shows that MMR can be usefully applied to vocabulary research and that qualitative research taking learners' perspectives into account is a useful way of complementing quantitative studies focusing on vocabulary measurement. The results of the qualitative strand indicate that intermediate learners can reflect on and provide information about their vocabulary learning process; hence, learners' accounts present an interesting source of data that has not been sufficiently capitalized on in studies on L2 vocabulary acquisition to date. In lexical research, qualitative methods have mostly been used in single-learner case studies (e.g. Fitzpatrick 2012), in interviews to test lexical knowledge (e.g. Pellicer-Sánchez & Schmitt 2010; Schmitt 1998) or validate the results of vocabulary tests (e.g. Pellicer-Sánchez & Schmitt 2012), and in research on vocabulary learning strategies (e.g. Bytheway 2015; Moir & Nation 2002). However, as this study shows, qualitative data can also be used to gain a different perspective on L2 vocabulary acquisition through the eyes of the learners. Collecting interview data from groups of participants allowed to explore informal vocabulary learning taking place in messy real-world contexts (Rose & McKinley 2017). By giving learners a voice and taking their lay theories seriously, useful insights were gained on how learners cope with unknown vocabulary in authentic input, which factors influence strategic attention, and which strategies are actually made use of. This avenue for research is certainly worth exploring further in the area of informal vocabulary learning, but it could also be fruitfully transferred to other areas of research on vocabulary acquisition.

The usefulness of qualitative methods for lexical acquisition research is, however, not the only conceptual and methodological insight gained with regard to vocabulary in the present study. In fact, a major contribution lies in the exploration of scoring issues in relation to the vocabulary tests used and the proposal of new and innovative scoring methods for both measures. As discussed extensively in section 5.3.3.3, V_YesNo (Meara 2015a) and Lex30 (Meara & Fitzpatrick 2000) offer several practical advantages, in particular easy and fast administration, which was a necessity in the case of the present study. However, as most other measures of vocabulary size (see section 3.2.3), they are also subject to limitations, especially with regard to scoring and the interpretation of scores (see also Beeckmans et al. 2001; Eyckmans 2004; Kremmel 2017; Meara 2009; Meara & Miralpeix 2017; Pellicer-Sánchez & Schmitt 2012; Walters 2012). Beginning with V_YesNo, issues as well as solutions developed in the present study are discussed below.

Since their introduction as fast and practical measures (Meara 2010; Meara & Miralpeix 2017) of receptive L2 vocabulary size, Yes/No tests have faced criticism with regard to scoring and over the years a number of different methods of correcting raw scores for overestimation of knowledge have been proposed (e.g. Beeckmans et al. 2001; Huibregtse, Admiraal & Meara 2002,

see sections 5.3.3.3 and 5.3.5.3). Recently, Meara and Miralpeix (2017) presented another innovative correction formula which uses an S-shaped logistic weighting function and thus "avoids excessive penalisation" of guessing (Meara & Miralpeix 2017: 119), by weighing the number of false alarms in relation to test takers' correct responses. This study used this newly proposed formula in combination with V_YesNo and is the first to compare the S-shaped logistic weighting function to previously suggested correction formulae such as *h-f, cfg,* Δm, and I_{SDT} (see Table B.10 and Figure B.2 in Appendix B). The analysis showed that V_YesNo scores calculated using the new logistic weighting function are reasonably close to the much simpler formula *h-f* (see section 6.4.1). If this similarity should also be found in future studies, it might be more efficient to use *h-f,* which has also been shown to be a good approximation in previous research (e.g. Pellicer-Sánchez & Schmitt 2012; Stubbe 2013).

However, going one step further this project also addressed the issue of lack of evidence in relation to the Yes/No format similar to previous efforts by Eyckmans et al. (2007) or Pellicer-Sánchez and Schmitt (2012). After completing V_YesNo, test takers were asked to either provide L1 translations or L2 explanations for 20 items, i.e. one fifth of the target words, used in the V_YesNo test. These additional items thus provide evidence of participants' actual meaning recall knowledge of the target words and can be used to assess the accuracy of their judgements on the YesNo test. Although originally these additional items were included to deter students from guessing, they can also be used to score test takers' responses on the V_YesNo test: first, participants' responses on the 20 translation items were compared to their judgements of the corresponding target words in the V_YesNo test to establish the proportion of correct judgements. This proportion includes all instances in which the judgement on the V_YesNo test matches the knowledge a participant demonstrates in the translation items (see section 5.3.5.3). In a next step, each test takers' individual proportion of correct judgements on the 20 selected items was multiplied with their overall number of correct responses on V_YesNo, thus generalizing from one fifth to all target words. The novelty of this score, which I named *hits corrected by the proportion of correct judgements* (*h×CJ%*), is that it takes evidence for participants' actual meaning recall knowledge into account. A comparison of the *h×CJ%* scores to the V_YesNo scores computed with the help of the logistic weighting formula indicates that the latter overestimates vocabulary size by over 900 words on average. This finding also has conceptual implications because it suggests that participants overestimated their receptive vocabulary knowledge on the V_YesNo test despite the relatively low number of false alarms, the exclusion of unreliable tests and the application of a correction formula (see section 5.3.5.3). The fundamental problem here seems to be that what L2 learners think they know does not necessarily correspond to what they actually know, at least not with regard to recall knowledge. Consequently, the results of the present study can be taken as further evidence that the construct of YesNo tests should probably be regarded as form recognition rather than meaning recall (see section 3.2.3). At the same time, it is crucial to point out that the newly proposed scoring method of *h×CJ%* is exploratory as well, hence, the truth might well lie somewhere in the middle.

The second vocabulary measure used in this study, Lex30, is usually scored with regard to frequency to gauge productive vocabulary size. Several issues emerged with regard to this frequency-based scoring procedure, which to an extent have also been noted in other studies: first, there is the problem of scoring cue words given as response words which results from a change in the frequency list used for scoring (Jiménez-Catalán & Moreno Espinosa 2005); second, the question of how to treat multiword expressions (e.g. Meara 2009) remains unsolved and third, the extent to which the accuracy of responses, i.e. spelling, should be taken into account is still unclear (see section 5.3.5.4).

The first issue resulting from changing the list recommended for scoring from Nation's (1984) list to the JACET 8000 list (Uemura & Ishikawa 2004) does not only concern cue words given as responses, as previously noted, but also extends to lemmatization. Meara and Fitzpatrick's (2000) original lemmatization criteria, which built on Bauer and Nation (1993), do not correspond to the criteria used for the JACET 8000 list. This creates additional problems because some response words produced in the present study would have scored a point only after the application of the lemmatization criteria (see the example in section 5.3.5.4). The present study addressed this issue by analysing the JACET 8000 list in relation to the proposed lemmatization criteria and developing a comprehensive scoring protocol to avoid distorting the scores in any way. The second issue of multiword units was dealt with as suggested by Meara (2009), but remains unsatisfactory. It is now widely acknowledged that formulaic language is an essential part of lexical knowledge (Schmitt 2004; Siyanova-Chanturia & Pellicer-Sánchez 2019; Wray 2002), but until frequency lists taking both single-word and multi-word expressions into account become available (see section 3.1.1), it will be extremely difficult to systematically award points to formulaic sequences produced in the Lex30 test. The third issue concerning accuracy of response words was explored by calculating two Lex30 scores: the regular score awards points regardless of spelling, whereas the adjusted Lex30 score subtracts half a point for each wrongly spelled word. Further analysis indicated that there is no relationship between the (unadjusted) Lex30 score and the number of mistakes made on the test, hence, only the unadjusted score was used for statistical analysis (see section 5.3.5.4).

Despite every effort to make the scoring procedure as objective and transparent as possible, the main issue with regard to scoring Lex30, i.e. that the final score resulting from the frequency-based analysis cannot be interpreted in terms of a concrete size estimate (Meara 2009; Walters 2012), could not be tackled in the present study. Instead, I have proposed a new and innovative scoring method for Lex30 data in addition to the frequency-based analysis. The schoolbook analysis (see section 6.4.7) was inspired by methodological comments made in Fitzpatrick and Clenton (2010) and the realization that in order to obtain a full picture of vocabulary learning from EE activities, formal educational contexts also need to be taken into account because they exert an influence on learners' informal engagement with the language in their leisure time (see also section 8.2). While the relationship between in- and out-of-school English learning was more centrally discussed in the focus group interviews (see section 7.5), the exploratory schoolbook

analysis presents an initial attempt to distinguish between productive vocabulary knowledge likely learned inside and outside the classroom and to draw tentative conclusions about the nature of vocabulary potentially acquired through engagement with EE.

Furthermore, the schoolbook analysis presents an additional possibility to examine the relationship between EE and productive vocabulary knowledge, which is especially interesting as the statistical model indicates that there is no significant relation between frequency of engagement with EE and productive vocabulary size as measured by Lex30. In this regard, it has to be pointed out that the multiple regression model for productive vocabulary size only explains 8% of the variance in Lex30 scores, while the model for receptive vocabulary size explains 21% of the total variance in V_YesNo scores. Although it is clear that many factors influencing lexical knowledge could not be measured in this study, this result appears to suggest that there are further variables affecting productive vocabulary size which are missing from the present research design and that these may be different from those influencing receptive vocabulary knowledge. However, this outcome could also be an indication that Lex30 is not sensitive enough as a measure of productive vocabulary size to show effects in relation to engagement with EE. This view is supported by the results of the exploratory schoolbook analysis: by comparing the responses produced in the Lex30 test to lemmatized lists of the vocabulary presented in the coursebooks used by participants at lower and upper secondary level it was shown that a surprisingly large amount of responses produced was not included in the schoolbooks and thus potentially acquired outside school (see section 8.2.2 for further discussion). While the nature of this additional analysis of the Lex30 data clearly is highly exploratory, it does present an interesting avenue for research that should be explored further.

In sum, the conceptual and methodological contributions of this thesis are located in three areas: first, a close investigation of concepts currently used in research on informal out-of-school language learning shows that extramural English is the most suitable conceptual approach for the present study because firstly, it incorporates both online and more traditional offline learning spaces and activities and secondly, it does not focus exclusively on learning but is interested in all learner-driven language practices taking place outside formal education. The empirical results support the use of EE as the central theoretical concept in that they show that Viennese adolescents engage with English in both on- and offline context and that language learning is not the primary driving force behind such informal English activities.

A second insight of this thesis concerns the methodological approach to researching difficult to access real-world phenomena such as EE. Mixing quantitative and qualitative methods in a fully integrated MMR design allowed to collect data on the frequency as well as on the amount of time spent with EE, participants' receptive and productive vocabulary size, several potential influencing factors, and teenagers' perspectives on engagement with EE and vocabulary learning. Furthermore, the use of a MMR design was valuable because the results of the two research strands could be used to confirm, explain and enhance each other, which led to a much

richer and more detailed perspective on Viennese students' EE activities and their relation to vocabulary knowledge.

Furthermore, the usefulness of MMR also extends to the third area of conceptual and methodological insights concerned with issues in vocabulary measurement. The present study did not only present new ways of scoring and analysing the data collected with the help of V_YesNo and Lex30 by introducing the *h×CJ%* score based on a comparison of translation items to Yes/No data and proposing an exploratory analysis method of Lex30 data in relation to participants' schoolbooks, it also showed that qualitative research taking learners' perspective into account is a useful way of complementing quantitative studies focusing on vocabulary measurement.

While the methodological suggestions made clearly need to be put to the test in further studies, the insights gained in this project suggest that a well-designed MMR approach based on a carefully considered conceptual foundation and the willingness to explore further innovative ideas can lead to rich empirical findings, which are summarized and discussed in the next section.

8.2   Integration and discussion of empirical findings

This section brings together the results of the two strands of this MMR project and presents empirical insights and meta-inferences based on both quantitative and qualitative findings. It is structured according to the three main research aims of the present study set out in section 5.1, namely, to map the landscape of extramural English among Viennese upper secondary school students (section 8.2.1), to explore the relationship between engagement with EE and receptive and productive vocabulary size (section 8.2.2), and to describe learners' perspectives on EE and their beliefs about learning from it (section 8.2.3). However, in keeping with the fully integrated MMR nature of this study, some results relating to learners' perceptions (aim 3) are discussed together with aims 1 and 2 since the qualitative and quantitative analyses mutually inform and influence each other. In addition, the integrated findings are contextualized and discussed in light of the literature presented in Chapters 2, 3 and 4.

8.2.1 Engagement with extramural English among Viennese teenagers

This section collates findings on the types and amount of contact with EE among students in academic upper secondary schools. Results show that participants in this study mainly engage in EE activities that involve popular media, such as music, films, series or video clips, and social media. Most of these activities involve electronic devices and/or a screen, many are carried out in online environments, and the vast majority of the most frequent activities involve language only in a receptive way. At the same time, participants' EE environments are characterized by diverse interests and individual preferences, which means that many teenagers also engage in seemingly 'unimportant', infrequent activities or specialized niche activities. In addition, data from the online language diary show that participants spend a large amount of their leisure time with EE with a mean of approximately four hours per day. This estimate is, however, credible in

relation to other studies and supported by students' statements in the focus groups. In the following, these findings are presented and discussed in greater detail together with information on the reasons for the use of English in extramural contexts.

As shown by both the quantitative and qualitative data (see sections 6.2 and 7.1), English is an integral part of Viennese teenagers' lives: 96% report encountering English in at least one leisure activity per day. Moreover, participants see English as the most important language next to German and other L1s; indeed, its knowledge and use appear to have become normal for many adolescents with some describing it as 'ordinary'. This is in line with the analysis of the wider Austrian context in Chapter 4, in which English emerges as the most important language next to German in a situation of 'globalized bilingualism' (Smit 2004; Smit & Schwarz 2020).

Further analysis points to two main reasons for these views (see section 7.1): the importance of English for leisure activities and its role as the global lingua franca. Clearly, the global position of English has an impact on the perceived importance of English for worldwide communication and international travel, and knowledge of English is thus seen as essential for participants' futures, which is also reflected in the quantitative data (section 6.2). The fact that currently English is the 'universal language', as one participant put it, also plays a role in participants' daily lives because it is seen as the default language of communication with non-German speakers and therefore used with tourists as well as with international family and friends. However, the reason that was most frequently given for the importance of English is the fact that young Austrians need it for their leisure time activities. While the availability of resources and media in English is evidently linked to its position as the 'universal language', this finding highlights the role of English in teenagers' spare time activities and shows that EE contributes to the perceived importance of English.

In previous research similar beliefs have been voiced by teenagers in Iceland (Ingvarsdóttir & Jóhannsdóttir 2017; Jeeves 2017), Norway and Poland (Anioł 2011), Sweden (Sundqvist 2019) and Spain (Nightingale 2016). Swedish teenagers in Sundqvist's (2019: 104) study also emphasize that English has become completely normalized in their lives and Norwegian and Polish adolescents in Anioł's (2011: 119) comparative study both "stressed that it is very difficult even to imagine not knowing English". In her study the importance of English was directly linked to media consumption for Norwegian but not Polish participants, and similar media-related explanations for the importance of English have also been given by secondary school students in Iceland (Jeeves 2017) and Spain (Nightingale 2016). Hence, like the present project, all of these studies point to an impact of informal leisure time activities on the perceived significance of English in adolescents' daily lives.

Yet, in addition to leisure-related reasons, there are also strong utilitarian motives for the perceived significance of English among young Europeans. The notion that knowing English is useful because it is widely used internationally is even present among young language learners between the ages of 8 and 12 as reported by Muñoz (2014). She found that this view is voiced more frequently with increasing age, and that the importance of English skills for future jobs is

already mentioned in this young age group. Likewise, using English abroad and for studies and/or jobs are the two most frequently named expected future uses among Icelandic students described by Ingvarsdóttir and Jóhannsdóttir (2017), followed by informal leisure activities. Nightingale (2016) also found that great usefulness is attributed to English and that Spanish participants thought that everyone, including future children, should learn it. Such views are also found in the Austrian data from the Special Eurobarometer on languages (European Commission 2012a), in which 93% stated that English was the most useful language to learn for children (see also section 4.2). Hence, the two main explanations for the importance of English found in the present study also emerged in other European research projects, which points to a certain level of comparability of European teenagers' views on the usefulness of English and its role in their lives.

As the first larger-scale investigation of Austrian, or more specifically Viennese, secondary school students' engagement with extramural English, this study further shows that they are comparable to their European counterparts surveyed in previous research (e.g. Ingvarsdóttir & Jóhannsdóttir 2017; Peters 2018; Olsson 2012; Sundqvist 2009a) also in terms of their exposure to extramural English. On average, learners in this study engage in 9.43 daily EE activities, which suggests that they come in contact with English through a variety of sources. Results concerning the types of extramural contacts show that there are some very common and widely popular EE activities, but that at the same time participants' EE environments are diverse and highly individual (see sections 6.3.1 and 7.2.1).

The top three activities in which more than 50% of participants engage on an (almost) daily basis and over 75% at least a few times a week are listening to music, watching video clips, and reading in social media. These activities coincide with students' preferred leisure activities overall (section 6.1.3) and thus one first conclusion to be drawn is that what Viennese teenagers like doing in general, they also do a lot in English. Other frequent activities in which over 50% engage at least a few times a week are singing, in particular singing along to music; watching films and series, mostly online; using apps and search engines; and using English words while speaking or writing in other languages. These highly frequent EE activities listed above fall into three broad categories: first, participants engage with English-language music through listening and singing; second, they watch 'online TV' in the form of English-language video clips, movies and series; and third, they read or use English in other online environments such as social media, apps or search engines. This categorization highlights three further conclusions about Viennese teenagers' modes of practice (see section 2.1): popular EE activities among Viennese teenagers typically involve an electronic device and often a screen, most are carried out online, and almost all only involve language in a receptive way.

As pointed out above, these results concerning the most frequent EE activities are similar to those of other studies focusing on European adolescents (Ingvarsdóttir & Jóhannsdóttir 2017; Lyrigkou 2018; Nightingale 2016; Peters 2018; Olsson 2012; Sundqvist 2009a), but also to teenagers' practices in very different parts of the world such as China (Lai, Zhu & Gong 2015),

Indonesia (Lamb 2004b) or Japan (Barbee 2013). These studies also found that out-of-school exposure to English revolves around media such as music, TV programmes, films, series, video clips and games, although the exact order of popularity varies. Similarly, research has shown that much informal contact with English happens online (Kusyk 2017; Lee & Dressman 2018; Sockett 2014; Sockett & Toffoli 2012). In the Viennese context, Hahn (2017), which is an approximate replication of the quantitative strand of the present study with students from vocational business middle schools (see section 3.3.1), found almost the exact same ranking of EE activities with the top three activities being identical. The finding that Austrian teenagers frequently engage in the same activities as their counterparts in subtitling countries such as Belgium or Sweden shows that traditional TV only plays a minor role in young people's media environments these days because they prefer more flexible on-demand offers via streaming. Consequently, one finding of this study is that the difference between subtitling and non-subtitling countries loses its relevance with regard to EE because the results on teenagers' practices are strikingly similar across countries.

However, in both Viennese studies one activity is conspicuous in its absence among the most frequent activities: gaming. While research in other contexts has found gaming to be a popular EE activity (De Wilde, Brysbaert & Eyckmans 2019; De Wilde & Eyckmans 2017; Hannibal Jensen 2017; Jóhannsdóttir 2017; Persson & Prins 2012; Sundqvist 2009a), playing digital games is not a frequent activity in Hahn (2017) and in the present study. However, that is not to say that there are no gamers in the sample; rather, it appears that while gaming is not a common activity among Viennese 10th-grade students with only 27% reporting to play English-language games on a computer or console at least a few times a week, those who do play games do so very frequently and often for extended amounts of time. Both the Upper Austrian youth media study (Education Group GmbH 2017a) and the Austrian media analysis (Verein Arbeitsgemeinschaft Media-Analysen 2017) also show that gaming is not among the top five media activities for young Austrians, although in these larger studies roughly 60% report playing digital games. Therefore, the absence of gaming as a frequent EE activity in this sample is surprising in comparison to international but not Austrian research.

Regarding the amount of contact with EE, results show that Viennese adolescents spend a surprisingly large amount of time with English during their spare time. Next to German, English is the language teenagers in the present sample use most for leisure activities, for 51% it is the second most frequently used language and for 34% the third most frequent language after a second home language (see section 6.3). More detailed data on time spent with EE were collected with the help of the Extramural English Online Language Diary. These show that 10th-grade students spend an average of 4 hours and 7 minutes with EE per day (see section 6.3.2). While at first this may seem to be an incredibly large amount of time, this finding is plausible for several reasons. First, analyses of the EEOLD data show that participants spend more time with EE on weekends than on weekdays and an examination in relation to participants' evaluation of their EE time as more or less than usual also supports the result. Second, the mean EE time of just over

4 hours concurs approximately with students' reports on time spent online (3 hours and 57 minutes, see section 6.1.3) and since much contact with English occurs in online environments such a relation appears reasonable. In addition, other Austrian studies have found comparable results concerning time spent online (4 hours and 44 minutes, IAB Austria, BVDW & IAB Switzerland 2016) and have shown that Austrian adolescents generally have a large amount of leisure time at their disposal (5 hours and 58 minutes, Statistik Austria 2010). Third, the estimate is in line with previous studies on Swedish teenagers' use of EE because a comparison of their findings indicates that the amount of time spent with EE is on the rise: Sundqvist (2009a) found a daily mean time of 2 hours and 38 minutes, in Olsson (2012) it was 2 hours and 54 minutes and in Olsson and Sylvén (2015) non-CLIL students spent 5 hours and 36 minutes with EE per day. Fourth, and most importantly, the quantitative outcome is also supported by the findings of the qualitative strand: most participants in the focus groups agree that the result is plausible upon consideration of their own EE habits (see section 7.2.2). In addition, they add important insights: participants explain that engagement with EE is not always entirely conscious, especially music is often used as a background to other tasks. This qualitative finding inspired further quantitative analysis, but even if all music-related variables are removed from the language diary data, mean EE time still amounts to 3 hours and 21 minutes. Students further report that sometimes EE activities are carried out simultaneously and, most importantly, a lot of EE activities appear to happen 'on the move', with teenagers using their smartphones to access English-language content or music while walking or travelling on public transport. These explanations correspond with theories of mobile-assisted language use (Jarvis & Achilleos 2013) and with previous research on informal language learning via smartphones (Jurkovič 2019; Lai & Zheng 2018).

So far, we have seen that most Viennese adolescents participating in the present study come in contact with English through very popular activities and that they spend a large amount of their leisure time with English. However, while such generalizations about the majority of participants are certainly useful, they tend to mask the diversity present in the sample. In fact, both quantitative and qualitative data show that Viennese 10th-grade students have highly individualized EE environments in accordance with their personal preferences (Berns, De Bot & Hasebrink 2007; Livingstone, d'Haenens & Hasebrink 2001). First, all 64 EE activities listed in the EEQ, even the least popular ones, are carried out at least a few times a month by at least one participant (see section 6.3.1). This is testament to the high diversity in the sample itself, but results show that in addition to these common activities many participants also engage in specialized 'niche activities', such as programming, rapping, sports or creative writing. While this study is not the first to highlight the diversity of learner's EE environments (e.g. Olsson 2012) and some research points to more beneficial effects of more varied EE activities on language development (Lee 2019a), it is nevertheless important to remind ourselves that individual students also engage in seemingly marginal activities.

Despite the strong influence of individual interests on Viennese adolescents' EE environments, some broader tendencies also emerge in the quantitative analysis (see section 6.3.4). Similar to previous studies (Hahn 2017; Olsson 2012; Olsson & Sylvén 2015; Peters et al. 2019; Sundqvist 2009a), the popularity of EE activities differs according to gender and there is a significant gender difference with regard to time spent on EE, with boys spending more time with English in out-of-school contexts than girls. Concerning types of activities, male participants engage significantly more often in all types of digital games and related activities such as in-game chats. Statistically significant differences in favour of boys were also found for watching online video clips and movies without subtitles, and reading comics. Female participants, on the other hand, engage significantly more often with music through listening, singing and reading lyrics and they read more books, information print texts and stories. They are also significantly more likely to write stories or a diary in English themselves. In addition, girls use social media in English more often with regard to reading and writing status updates, but there is no significant difference with regard to messaging.

Studies in other contexts have also found large gender differences with regard to gaming (Hannibal Jensen 2017; Olsson 2012; Peters et al. 2019; Sundqvist 2009a) and that males generally engage in more EE activities; however, in Olsson and Sylvén (2015) the overall gender difference is explained solely by the discrepancy in gaming, there is no difference with regard to other activities. While gaming thus seems to be the biggest gender divide, this study also found other differences: similar to Olsson's (2012) findings, girls read more often and engage more extensively with music and some aspects of social media, whereas boys play games and watch video clips more frequently like in Hahn's (2017) study.

Regarding further influencing factors, socioeconomic status and overall English proficiency show positive correlations with the EE median score in the bivariate analysis (see section 6.3.4). The effect size for the correlation with SES is, however, small and since Sundqvist (2009a), one of the few other studies to take SES into account, found no statistically significant effect, it is difficult to draw conclusions. English proficiency as measured by an CEFR-based self-assessment tool correlates with EE with a medium effect, but surprisingly there is no relation between EE and length of English instruction. Hence, a tentative conclusion could be that for frequency of engagement with EE it is not important how long students have been studying English, but which level of proficiency they have achieved during this time. In relation to the question of what comes first – more EE or higher English proficiency – this finding could point to a two-way relationship: more proficient students probably engage in EE activities earlier and more frequently than their peers, and this in turn may lead to practice and learning effects and thus to an increase in proficiency. However, at this point we can only speculate about the evolving relationship between EE and language proficiency; further research on early engagement with English and longitudinal studies tracking developments over time are needed to provide more definite conclusions (see also section 9.3).

With regard to the relation between EE and language development it is also informative to compare the use of language skills in informal environments. As mentioned above, the most popular EE activities identified are related almost exclusively to receptive language skills and the same is true for EE activities done a few times a week or a month, although students are more likely to encounter written English in these (see section 6.3.1). Data on the amount of time spent with English show a similar trend with more than two hours a day spent on listening, almost one hour on reading and half an hour or less on speaking, writing and multi-skill activities (see section 6.3.2). These results on the use of the fours skills in EE activities are comparable to the findings of the previous studies carried out in Austria (Hahn 2017; Miglbauer 2017; Wieland 2016) and other contexts such as Belgium (Peters 2018), Iceland (Jóhannsdóttir 2017), or Sweden (Olsson 2012; Olsson & Sylvén 2015; Sundqvist 2009a; Sylvén 2004/2010). However, while all of these studies have found a predominance of receptive skills across contexts and age groups, it is important to highlight that less frequent niche activities often involve much more language production and that half an hour of speaking and/or writing in out-of-school contexts probably still amounts to more productive language use outside than inside the classroom.

Participants also provide reasons for the lower amount of productive language use, mainly relating to the amount of effort required and the absence of a need to communicate in English. Writing in English, especially in online contexts is seen as strenuous and only participants with multilingual friend groups see a need to use English on social media, whereas participants with a majority of German-speaking friends do not use English for public statements on social media (see also Lai & Zheng 2018), although a few report using it for messaging with Austrian friends. Speaking English appears to be even more limited and is restricted to gaming, multilingual peer groups, and niche activities such as English-speaking sports teams.

Data on the more general reasons for teenagers' use of English in extramural contexts are more informative and primarily point to a conscious decision to use English for aesthetic reasons and to the greater and earlier availability of content in English as a second factor (see sections 6.3.3 and 7.2.3). In both datasets the two top reasons for using English are the fact that it is seen as cool, beautiful or sounding better and that participants prefer original versions, mostly with regard to films and series but in some cases also concerning games and books. Besides, availability also plays a role, but it is not the most decisive factor: participants state that a lot of online content or music just is in English and some mention that being up-to-date, which means watching or reading original versions rather than waiting for dubbed or translated versions, also plays a role. In addition, participants argue that English provides access to a wider pool of information and in some students' views better (online) content is available in English. These views on the quality of information are similar to those expressed by adolescents in Anioł (2011), who also found a preference for original versions among Norwegian but not Polish teenagers. With regard to dubbing, a large-scale survey conducted for the European commission with over 5,000 participants indicates that young, multilingual Europeans generally prefer watching

original versions with subtitles: "the younger the respondents (aged 12-18 and 18-25) and the more languages they speak, the more pronounced is their preference for subtitling over dubbing" (Media Consulting Group/EACEA 2009: 10).

In addition, almost 80% of the students also agree with the statement that improving their English skills is another factor for engagement with EE in the quantitative data, while at the same time nearly as many (73%) say that they just enjoy using English. These results point to the conclusion that participants see developing their English skills as a welcome by-product of using English in their spare time, which is also supported by qualitative findings on students' perceptions of learning from EE discussed in section 8.2.3. Similarly, a study with university students in Hong Kong (Lai 2015) also found that language use in out-of-school English activities is casual and spontaneous and that any learning is non-systematic and not the primary purpose. These findings highlight that learning is usually not the main aim of using English for leisure activities and support the position adopted in Chapter 2 that extramural English is the most appropriate concept because it does not make any assumptions about the purpose for engaging with English (see sections 2.2 and 8.1). The use of English outside of formal education does, however, certainly provide opportunities for lifewide language learning (Banks et al. 2007) and the next section will discuss in how far these are related to vocabulary acquisition.

8.2.2 The relation between extramural English and vocabulary knowledge

This section is centrally concerned with the present study's findings on vocabulary knowledge and its relationship to extramural English. The results indicate that EE affects receptive and productive vocabulary knowledge differently: for receptive vocabulary size the quantitative analysis shows a positive and statistically significant relationship with EE, whereas neither the bivariate nor the multivariate analysis shows a relationship between EE and productive vocabulary size. At the same time, the results of the exploratory schoolbook analysis and the fact that participants in the focus groups were able to provide examples of lexical learning from EE indicate that the acquisition of some productive vocabulary knowledge is possible in extramural contexts, although to a much lesser extent than for receptive knowledge. Concerning other factors influencing vocabulary knowledge, length of instruction emerged as a significant predictor of both types of knowledge in this study, but for receptive vocabulary slightly more variance was explained by the EE predictor. In the bivariate analysis only, the SES summary index and the number of books present at students' homes also show small significant positive correlations with both receptive and productive vocabulary size. Although boys achieved higher scores on both tests, the gender difference was statistically significant only for productive vocabulary size. Below, the results on receptive and productive vocabulary size, extramural English and other influencing factors are again presented in more detail and contextualized in light of previous studies.

Before exploring relations between vocabulary knowledge and extramural English further, it is informative to compare the findings on vocabulary size in this study to earlier research in the Austrian and European context. Results of the V_YesNo test indicate that the mean receptive

vocabulary size of Viennese 10th-grade students is 4,807 lemmas according to the logistic weighting function (V_YesNo score) and 3,847 lemmas according to hits adjusted by the proportion of correct judgements (*h×CJ%*). This result is considerably smaller than the 7,690 word families found for Austrian 10th-grade learners in Zichtl (2017), but as discussed in section 3.1.2, this very large result appears to be at least partially explained by a strong facilitative effect of German-English cognates on the VST. Henriksen's (2008) findings with Danish students in grade 10, the majority of which had mastered the 2,000 most frequent word families, but not the 3K level, seem more plausible in comparison. In the second Viennese study, Hahn (2017) found a lower receptive vocabulary size of 3,041 lemmas for 10th-grade students in vocational middle schools, using V_YesNo in combination with the logarithmic weighting function. The lower result found in her study is consistent with curriculum specifications: participants in Hahn's study are supposed to have reached CEFR level A2 in comparison to learners in this study who should have attained level B1 in grade 10.

Concerning productive vocabulary size, the mean Lex30 score among the participants of the present study is 38.23 with a range of 7 to 69 points and a median of 37. Walters (2012) found comparable results for an intermediate group of Turkish university students, who had a mean score of 36.72, but unfortunately, no further indication of proficiency level is reported in her study. In comparison, Fitzpatrick and Clenton (2010) found a lower result among 40 Japanese medicine students with a mean of 24.3 points. Investigating Spanish students in CLIL contexts, Alejo González and Piquer Píriz (2016) administered Lex30 to two groups of 14- to 15-year-old students in the third year of secondary school, who had received approximately 900 hours of English instruction. The mean score in group A was 33.37 and the mean score in group B was 29.48. These results are slightly lower than those of the present study with 15- to 16-year-old learners who received approximately 805 English lessons amounting to 671 full hours. Clearly, further investigation to aid the interpretation of Lex30 scores is needed (see sections 3.2.3 and 5.3.3.3) but overall these comparisons with other studies indicate that the receptive and productive vocabulary sizes of the participants in this study are comparable to earlier studies in similar contexts and therefore not implausible despite the measurement issues discussed extensively in sections 5.3.3.3 and 8.1.

The central question of this study regards the relationship between vocabulary size and extramural English. In the multivariate analysis which estimates the effects of EE, length of English instruction, SES, media access, gender and the number of home languages on vocabulary size through standard multiple regression analysis two variables emerged as statistically significant predictors for receptive vocabulary size: frequency of engagement with EE and length of English instruction (see section 6.4.3). In contrast, in the model for productive vocabulary size only length of instruction is a statistically significant predictor (see section 6.4.6). Hence, one of the key findings of this study is that the frequency of contact with extramural English shows a positive relationship with receptive vocabulary size, but not with productive vocabulary size. This result is entirely plausible because, as discussed in the previous section,

the EE activities that participants engage in most frequently and for the longest amount of time mainly involve the receptive language skills of listening and reading, whereas language production plays a subordinate role in out-of-school contexts. In the two models for receptive vocabulary size, frequency of engagement with EE is the strongest predictor explaining the largest amount of unique variance, even though that amount is very small overall. Notably, contact with English through leisure activities outside school explains more variance than length of English instruction in both the V_YesNo and the *h×CJ%* model.

Hence, while EE emerges as the strongest predictor of receptive vocabulary size among the predictors considered in the multiple regression model, the outcomes for productive vocabulary size seem to suggest that it is not affected by engagement with extramural English (see section 6.4.3). Even engagement with highly specific niche activities, which were hypothesized to be more beneficial for productive vocabulary acquisition because they tend to involve more productive language use than the most frequent activities, does not show any effects in relation to productive vocabulary size. Indeed, the 43 participants who report engaging in one of the eight least frequent EE activities listed in the EEQ or named additional ones have a significantly higher mean receptive but not productive vocabulary sizes than the remaining participants in the sample(see sections 6.4.2 and 6.4.5).

However, the conclusion that productive vocabulary knowledge is not affected by EE stands in contrast to the results of the exploratory analysis of the Lex30 data against participants' schoolbooks which shows that participants produced lexical items in the association task that did not come up in their coursebooks (see section 6.4.7). About one fifth of the responses were not found in the participants' respective schoolbooks, which makes it likely that at least some of these lexical items were acquired in extramural contexts, although clearly teachers may introduce vocabulary not included in the coursebook. The response words that are not included in schoolbooks mainly fall into the category of mid-frequency vocabulary and thus occur relatively frequently in authentic language use, which supports the conclusion that a large proportion of these off-list types are encountered outside school. Moreover, categories identified in a thematic analysis, such as vocabulary related to modern and historical warfare, death, fantasy, and pejorative terms, indicate that at least some were acquired from EE because such vocabulary is used more frequently in informal, out-of-school language input like games, series, stories or social media than in the more formal language use of educational contexts. Clearly, this analysis is very exploratory in nature, but it indicates that productive vocabulary knowledge may benefit from EE activities, even though such an effect has not been found in the statistical analyses. In addition, participants in the focus groups were able to provide examples of vocabulary learned in out-of-school contexts (see section 7.4.2), which again points to the conclusion that some lexical knowledge can be acquired at a productive level from EE. The examples also highlight the diversity of the focus group participants' EE environments and show that in addition to informal language and slang terms, learners can also acquire infrequent, formal or specialized vocabulary from engagement with EE.

Overall, these findings suggest that extramural English affects receptive and productive vocabulary knowledge differently, with the statistical models indicating that there is a stronger relationship with receptive vocabulary knowledge. This outcome is not entirely surprising because the most common EE activities mostly involve receptive language use and are therefore less likely to support the development of productive knowledge. Unfortunately, it is rather difficult to compare this key finding to other studies because few have investigated productive vocabulary knowledge in addition to receptive vocabulary size. Sundqvist (2009a) found a significant positive correlation with a medium effect between a vocabulary index variable combining the results of the VLT and PVLT and an EE index variable, but her description suggests that EE was also positively related to the separate results of the receptive and productive test. Hahn (2017) also found moderate correlations between the number of daily EE activities and receptive and productive vocabulary size as measured by V_YesNo and Lex30, but her analysis did not include multivariate modelling. Hence, further research is needed to clarify whether and in how far engagement with EE does indeed affect the two types of vocabulary knowledge differently.

Concerning the relationship between EE and receptive vocabulary size, the results of this study point to a significant, if rather small, correlation similar to previous research (Berns, De Bot & Hasebrink 2007; Peters 2018; Peters et al. 2019; Verspoor, De Bot & Van Rein 2011). The second significant predictor was length of instruction, which corresponds to effects found for educational level identified in several studies (Peters et al. 2019; Verspoor, De Bot & Van Rein 2011). Moreover, similar to Peters' (2018) study in Flanders, frequency of engagement with EE explained more variance in receptive vocabulary size than length of English instruction at school, although Peters (2018) found a much larger difference in the amounts of variance explained by the two variables than the present study. Comparable results in relation to further aspects of lexical knowledge have been identified by Muñoz (2011), Schmitt and Redwood (2011), and González Fernández and Schmitt (2015).

The remaining influencing factors did not emerge as significant predictors in the multiple regression models, but SES and the related variables of the number of books and the number of different media devices available at students' homes did show statistically significant relationships with receptive vocabulary size in the bivariate analysis (see section 6.4.2). The SES summary variable and the number of books also correlated significantly with productive vocabulary size (see section 6.4.5). Surprisingly, SES has only been taken into account in relatively few studies, which is also due to missing data because of selective non-response (e.g. Puimège & Peters 2019). Contrary to this study, Hahn (2017) found a small, but significant effect of SES on productive, but not on receptive vocabulary, whereas two studies focusing on learners in primary schools found no effect for SES (Persson & Prins 2012) or parental education (De Wilde & Eyckmans 2017). In contrast, gender differences, which in this study have only been found to be significant in relation to productive vocabulary knowledge although boys achieved higher scores on both measures, have been examined in a number of studies: at secondary level,

both Sundqvist (2009a) and Hahn (2017) found that boys had significantly larger receptive, but not productive vocabularies, whereas Peters (2018) and Peters et al. (2019) found no difference according to gender. At primary level, Puimège and Peters (2019) also found an advantage for boys, whereas De Wilde and Eyckmans' (2017) results showed no effect for gender. In sum, further research is needed on these two influencing factors because findings vary considerably, even if studies are conducted in very similar contexts, such as Hahn (2017) and the present project.

While this section focused on the quantitative findings regarding the relationship between EE and receptive and productive vocabulary knowledge, the next section presents and discusses complementary qualitative findings on learners' views of EE and language, or more specifically vocabulary, learning.

8.2.3   The learners' perspectives on extramural English and language learning

This section addresses students' perceptions of extramural English and language learning with a particular focus on vocabulary acquisition as well as their views on language learning practices inside and outside formal educational contexts. Findings show that the majority of participants believe that engagement with EE benefits their language development, while some express more mixed or even negative views. Students name a number of factors which either help or hinder language learning from EE in their view; an unquestioned belief in the benefits of exposure to native speakers becomes apparent in this regard, but factors such as repeated encounters, familiarization and practice effects, and a positive influence on affective variables such as motivation are also discussed. The main obstacles identified by the focus group participants include encountering incorrect language structures, often in relation to slang, comprehension problems, and wrong inferences. With regard to the learning potential of individual EE activities, students evaluate productive language use through speaking and writing as well as reading positively, although these generally are not frequent activities. Listening to music as the most popular activity received mixed to negative evaluations and the benefits of gaming largely depend on the type of game played, in participants' opinion.

With regard to learning effects, vocabulary emerges as the most frequently name object of learning from EE; however, the interview data also show that learners need a good reason to invest effort into discovering the meaning of an unknown lexical item during their spare time: salience and importance for understanding the content, repeated encounters, and formal properties that attract attention emerged as factors in this regard. The default strategy that students use when they come across new English lexical items they want to understand is guessing from context. In addition, some report thinking about other languages they know or using an online dictionary, but many only engage in such strategic behaviour if their standard inferencing strategy fails and the target item is important for understanding the context.

Finally, concerning the relation between in- and out-of-school English learning most participants in the focus groups do not see a link between their English practices inside and outside school,

but many agree that school lessons at lower secondary level provide the basics necessary for extramural language use and informal learning. At the same time, several participants argue that at their current level more is actually learned from using English outside than inside the classroom, although the transfer of extramurally acquired knowledge to the school context is sometimes seen as problematic. These insights into learners' views on EE and language learning are again discussed in more detail in the following.

As mentioned, most participants in the focus group interviews express positive views concerning the effects of EE on language development (see section 7.3.1): 22 out of 30 participants believe that they learn from engagement with EE. In fact, many of these learners state that EE activities helped considerably in developing their English proficiency and some even contend that it contributes more than lessons at school. In contrast, five participants expressed mixed views: these students either think that engaging in EE activities helps to practise their English skills but does not lead to the acquisition of new knowledge, or they explain that although one could learn from EE activities, they hardly ever engage in any themselves. Only two male students express negative views on learning from EE: both of them are gamers and they state that they mostly encounter 'bad English' outside school, which is more likely to hinder rather than support their language learning.

With regard to what helps or hinders learning from EE, students came up with a variety of factors (see section 7.3.3). In most focus groups there was a strong and unquestioned belief that exposure to or communication with native speakers (NS) of English helps learning and, conversely, the most frequently mentioned obstacle is learning 'wrong things' from people who speak 'bad English'. This belief is also the main reason why the two gamers mentioned above evaluate learning from EE negatively. While it may be true that the quality of input in informal settings is not the same as in educational settings, this unquestioned equation of 'good English' with L1 English is slightly troubling at a time when ELF speakers clearly outnumber native speakers (e.g. Crystal 2003; Graddol 2006). The perception that L1 varieties of English are somehow better than the varieties spoken by L2 users is still prevalent among many learners as shown in a review by Subtirelu (2013). In his own MMR study, Subtirelu (2013) found that participants' reports were often contradictory in themselves. This ambiguity appeared to stem from differences between ideal and pragmatic goals, different understandings of what it takes to fulfil NS norms, and tensions between learners' own plans and the perceived expectations of others. However, a study by Ke and Cahyani (2014) suggests that informal contact with English can have positive effects: their participants showed a strong NS bias but engaging in online communication with other L2 speakers had a positive effect on students' attitudes towards ELF and their own linguistic confidence.

In addition to the strong belief in authentic L1 input, participants in the present study also mention a number of other factors that support learning from EE in their view. One aspect that was mentioned by almost as many participants is repetition, referring to both repeated encounters with a language structure and repeated engagement in an EE activity. As discussed

below, this factor was mentioned specifically in relation to vocabulary acquisition, but learners also state that repeated engagement with EE activities over time has the additional benefit of leading to a familiarization effect, particularly in terms of ease of comprehension, and to practice effects because EE input triggers previous knowledge or because they use already existing skills. Furthermore, many participants argue that being able or even forced to use English productively during EE activities is more helpful for language development than receptive exposure to English. For some this belief is again linked to NS norms in that they only regard speaking to natives as effective, but others also connect the idea to international (online) communication without reference to NS norms. Out-of-school engagement is also seen as beneficial in relation to affective factors: participants in the focus groups emphasize that it enhances motivation because students are genuinely interested in these activities. Such a positive impact of EE activities on motivation has also been noted in other contexts, for instance among university students in France (Toffoli & Sockett 2013), Hong Kong (Lai 2015) and Japan (Barbee 2013); however, this increased motivation appears to be strongly linked to the enjoyability of EE activities (Barbee 2013) and does not necessarily extend to the classroom context (e.g. Sundqvist & Olin-Scheller 2013; Ushioda 2013).

Participants also identified further aspects that hinder learning in addition to the issue of 'bad English input' mentioned above. Interestingly, some of these relate specifically to English-language music because learners state that in songs language is different and that one may learn incorrect grammatical structures due to the use of slang, although this caveat clearly applies to other EE input as well, for instance to social media. Moreover, individual students also show awareness of issues relating to contextual learning, such as the danger of making wrong inferences, which is also found in the literature (e.g. Elgort 2017; Van Zeeland 2014). They also note that the English encountered outside school is sometimes too difficult to learn from, which relates to issues of comprehension and coverage discussed in section 3.1.2, and that one sees or hears the same structures over and over again so that unknown structures are not encountered frequently enough to acquire them, which relates to overall frequency effects in language learning (section 3.1.3). Overall, however, participants named more positive aspects and also discussed them more frequently in the interviews, which supports the finding that most of them do believe their language development benefits from engagement with EE.

In addition to the overall learning potential, participants also discussed individual EE activities (see section 7.3.4). As mentioned above, several interviewees in the focus groups agree that using English productively in face-to-face interaction or online communication is most beneficial to language learning. However, they expressed different views on which of the two contexts, i.e. face-to-face or online, and the two modes of communication, i.e. spoken or written, is more useful and sometimes provide contradictory explanations. This is a typical example of conflicting lay theories expressed in the focus groups, which become especially prominent during discussions of individual activities. Reading is also seen as helpful by a number of students, especially because they can control the speed of reading and reflect about unknown

language structures. In addition, some argue that in comparison to films and series they concentrate more on the language while reading. Yet, the participants also point out that unknown language has a greater impact when reading because there are no visuals to support comprehension. In comparison, fewer participants argue that audiovisual media provide the best learning opportunities, but those who do, stress the value of multimodal input for increased comprehension and retention as well as the fact that audiovisual media provide the opportunity to learn spoken word forms.

The learning potential of two further activities is rather contested in the focus groups and led to heated discussions: listening to music and gaming. The learning potential of music and songs is evaluated negatively by a number of students because of non-standard language use and a lack of concentration on the lyrics. However, others emphasize that if one pays attention to lyrics by looking them up, songs can potentially become a source of learning, particularly in relation to genres like rap and hip hop. Several participants also mention repetition effects in relation to music and state they learn frequently repeated catchphrases or whole lyrics if they listen to the same songs over and over again. Opinions diverge even more with regard to the learning potential of digital games. One factor that emerges as decisive is type of game: participants evaluate single-player games, in which language input is provided by the game only, as better than multiplayer games in terms of accuracy but they also note that in many games the same structures are constantly repeated. Concerning multiplayer games, several students agree that English-language interactions with other players offer little learning opportunity because the language is simple and often incorrect; others, however, argue that using English productively yourself is beneficial to language learning. Interestingly, despite the heated discussions on multiplayer games, one point was uncontested: the fact that the language used for communication with other non-German-speaking gamers is English (see also Bytheway 2015; Sylvén & Sundqvist 2012a). This implicit agreement again highlights the role of English as a lingua franca, which pervades most areas of life nowadays.

While to the best of my knowledge there are no other reports on learners' evaluations of individual EE activities in terms of their learning potential, interviewees' beliefs can be compared to empirical findings in other EE studies (see section 3.3.1). Reading was indeed found to have positive effects (Peters 2018; Sylvén 2004/2010) and interaction on social media and speaking can also support language development (De Wilde, Brysbaert & Eyckmans 2019). Regarding audiovisual media, studies (Peters, Heynen & Puimège 2016; Peters & Webb 2018) also show learning gains in terms of vocabulary. Unlike some participants in this study, research further supports the view that gaming is advantageous for language learning (De Wilde, Brysbaert & Eyckmans 2019; Peters et al. 2019; Sundqvist 2009a), whereas there is mixed evidence for music (Berns, De Bot & Hasebrink 2007; De Wilde, Brysbaert & Eyckmans 2019). Students' evaluations of the learning potential of these EE activities are thus surprisingly similar to previous research outcomes, although there is of course variation among the individual participants. Based on their previous experiences and beliefs, learners perceive different

affordances for learning in extramural contexts (see also Lai 2015), which is reflected in different points of view and conflicting lay theories presented in the focus group interviews.

In contrast to their differing perspectives on learning potentials, many participants agreed on what they learn most from engagement with EE activities: vocabulary was most frequently named followed by pronunciation. As discussed in section 7.3.2, this result may be biased because the students were aware of the vocabulary focus of the present study; however, other aspects mentioned spontaneously like learning idioms, casual or colloquial expressions, pronunciation or spelling also point to lexical learning. Furthermore, in previous research not focusing on vocabulary acquisition "[v]ocabulary emerged as the most prominent aspect learned by students of English out-of-school" (Kalaja et al. 2011: 52) as well. Besides aspects related to lexis, several participants referred to developing their procedural knowledge and a feeling for what 'sounds right'. Lai (2015: 274) reports a similar finding for university students in Hong Kong who also reported that using English outside class "gave them a stronger sense of the language in terms of how the language is actually used." Lastly, some learners in this study stated that they learn new grammatical structures, gain better comprehension or practise speaking freely.

In relation to vocabulary as the main focus of the study more detailed information on students' learning experiences and learning strategies was collected. In both quantitative and qualitative results the main strategies that are used when unknown English words are encountered in extramural contexts are guessing from context, thinking about other languages and consulting (online) dictionaries (see sections 6.1.1 and 7.4.1). Yet, in the focus groups opinions concerning dictionaries differ sharply: some participants see dictionaries as an easy and fast option to discover the meaning of an unknown word and regard it as less strenuous than thinking yourself, whereas others see it as the very last resort that is only used when all other strategies failed. Nonetheless, these three strategies are far more popular than all other discovery strategies; students do not frequently think about the part of speech or known word parts and they rarely ask others for help because more proficient language users are seldom in close proximity and contacting others via text or messenger simply takes too long in their opinion. In sum, while the options available to students to discover the meanings of unknown lexical items may have changed, these results are still rather similar to Schmitt's (1997) study, which found that dictionaries and guessing from context were the most frequent strategies.

In addition to the types of strategy used, the qualitative data show that only certain words are regarded as important enough to use strategic behaviour. Unknown English words that are considered irrelevant, aptly described as 'random talk' by one participant, are not worth looking up or reflecting upon. Hence, perceived importance determines the amount of strategic attention given to a new lexical item. In addition, guessing from context emerges as the default behaviour, further strategies are only applied if it fails or if a word is perceived as particularly important. In this regard, three factors are mentioned in the focus group interviews: first, perceived importance appears to depend on the importance of an unknown word for

understanding the content and on the amount on information provided by co-text and context. Only unknown words that are crucial for understanding the content *and* difficult to infer receive further strategic attention, for instance by looking them up in a dictionary. This relates to the salience of words and the availability of contextual cues, for which previous studies (e.g. Elgort & Warren 2014; Qian 2005; Van Zeeland 2014; Webb 2008) have found positive effects on vocabulary learning. Second, repeated encounters are a factor: if new words are met several times in EE input, they become worth looking up. Although the central role of frequency of exposure is a known factor in vocabulary learning research (see section 3.1.3), this relation between number of encounters and use of VLS is a new and interesting insight. Third, some participants report using strategies for words that somehow arouse their interest or attract their attention. Although participants' statements remain rather vague in this regard, some appear to be attracted by formal criteria such as spelling or pronunciation, whereas others report wanting to know the exact meaning of words that 'get stuck' in their head. This finding highlights that salience for a new lexical item cannot only be achieved through importance for comprehension, but also through formal properties (Ellis 1999).

Moreover, participants also mention a variety of factors which aid the retention of new words. Similar to research stressing the importance of frequency of occurrence (see section 3.1.3 and Webb (2014) for a review), students think that repeated encounters with words – both in extramural input and at school – help them remember new words. They also report that the salience of words in context and their importance for understanding the content plays a role, as does the use of strategies: looking up the exact meaning of a word or attempting to use it themselves supports retention in their opinion. Regarding lexical inferencing, there is disagreement on whether guessing words from context helps to remember them: some participants seem to argue that the mental effort invested helps to retain at least a vague meaning for these words, whereas others think that it is not enough to remember a new word, particularly if the goal is to know it well enough to be able to explain or translate it. While the first argument is similar to the Involvement Load Hypothesis (Laufer & Hulstijn 2001) in that it posits that deeper processing leads to higher retention, research has shown that lexical inferencing can but does not necessarily lead to learning (Elgort 2017; Pulido 2007; Wesche & Paribakht 2010). Yet, although students do not necessarily agree on what exactly helps them to retain new vocabulary encountered in informal activities, the majority clearly agree that their engagement with EE supports vocabulary learning overall.

Finally, the relation between language learning in in- and out-of-school contexts was a topic of interest in the focus groups as well. The interview data show that the majority of participants see no link between English lessons at school and their EE activities. The two contexts are regarded as separate spheres because what teenagers do outside school is or cannot be used at school and what is learned in lessons is considered irrelevant for their informal spare time activities. Similar views were found among secondary school students in Germany (Grau 2009) and Finland (Ranta 2010) and even among 11-year-old pupils in Sweden (Bunting & Lindström

2013). However, a few interviewees in this study argue that there obviously is a two-way relationship between informal activities and English lessons and that they mutually influence each other in terms of learning and practice effects.

A third viewpoint is that English teaching at school forms the basis for further language learning from engagement with EE, a perspective which is also supported by some students who see a mutual influence. Proponents of this third view argue that the basic knowledge acquired in the first years of secondary education provides the foundation for further language learning and use outside school and assert that their (early) EE activities would not have been possible without this preparation. In particular, basic grammatical structures could not have been acquired through informal leisure contact only in their view. Similar beliefs which portray classrooms as the place for learning the basics have also been identified in other studies: Lai (2015: 273) found that classroom study was perceived as necessary "to build the foundation" by university students in Hong Kong and in a study comparing Swedish and English learning in Finland Kalaja et al. (2011) showed that school is the place where the fundamentals are learned for both languages, but for students of English learning was not limited to schools.

At the same time, several participants agree that by now they learn more from their independent language activities than from school lessons. This view goes hand in hand with the belief that the basics are, and to an extent need to be, learned at school; these participants simply argue that once the foundation had been laid, they started building on it in extramural contexts. In fact, several students believe that overall school is only responsible for the smaller part of their English proficiency, as already mentioned. Comparable results were found in Sweden and Iceland: Henry (2014) reports that in a large-scale survey more than half of students stated that they learn as much or more English outside school than in lessons and in Jeeves' (2017) study some students even went as far as saying that they would not need English lessons at all because in their view they do not learn anything. One reason for this perception may be that teenagers see their informal learning from EE activities as more relevant: when 'practising English in the free world', as one participant put it, they can make choices on what language content to engage with, whereas the content of their English lessons is often regarded as uninteresting and repetitive. Another factor contributing to the perception that more is learned outside school may be the difference in amount of contact: almost 70% of the participants in the quantitative strand report using English more in their leisure time than in school lessons and the mean EE time of roughly four hours a *day* exceeds the three hours of school lessons a *week* by far. A few participants in the focus groups even argue that engagement with English outside school is necessary to succeed in school because if the majority improves their English through informal activities those who do not are left behind. While this a strong claim, other students agree that engagement with EE supports learning and achievement in school. However, some learners also point to problematic aspects in transferring knowledge acquired from EE to the school context: in particular, they refer to using language structures acquired from informal discourse in class

and being corrected by teachers because these structures were considered context-inappropriate or incorrect.

In sum, the majority of participants in the focus groups do not see a link between their current English practices inside and outside school, but most agree that school lessons at lower secondary level provide the basics necessary for extramural language use and informal learning. This finding points to a crucial difference between Austria and other European research contexts: while the divide between 'school English' and 'real-world English' (e.g. Grau 2009; Ranta 2010) was also found in the present study, many Viennese students see school as a basis for engagement with EE. This stands in marked contrast to findings from European subtitling countries (e.g. De Wilde, Brysbaert & Eyckmans 2019; Kuppens 2010; Persson & Prins 2012), where many children begin to learn English from extramural activities before the start of formal instruction. Consequently, the results of this study indicate that the dubbing practices of Austrian media are not relevant for adolescents' engagement with EE (see section 8.2.1), but they do influence the learning trajectory of Austrian learners of English who, in contrast to children in subtitling countries such as Belgium or Sweden, build the foundations for engagement with EE through formal education at school. Hence, more research on links between in- and out-of-school research as well as on the implications of extramural activities on teaching practices is needed across different contexts and educational levels. Further ideas for future research directions are presented in Chapter 9 together with a summary of the study and its key findings.

9 Conclusion

This chapter concludes the thesis by summarizing the study and situating it in a wider context in relation to both research and teaching. First, section 9.1 recapitulates the rationale, main aims and design of the study before providing an overview of the content covered in each chapter. The study's key findings on extramural English and its relation to vocabulary knowledge and learning as well as their implications for English language teaching are synthesized in section 9.2. Finally, section 9.3 highlights the significance of the study for the Austrian context and beyond, while at the same time acknowledging its limitations. The chapter concludes with a look to the future and a discussion of possible directions for further research.

9.1 Summary

This study explored teenagers' English practices beyond the walls of their classrooms and examined the relationship between their EE activities and vocabulary knowledge. It is the first larger-scale project on engagement with extramural English among secondary school students in Austria, which presents a rather new research environment. In line with global tendencies, the presence of English in Austria has increased considerably over the last decades so that a wide range of opportunities for using as well as potentially learning English in out-of-school contexts are available these days. Focusing on 15- to 16-year-old teenagers attending academic secondary schools in the capital Vienna, the study pursued three main aims: first, it established an overview of extramural English practices by collecting data on the types and amount of contact with EE. Second, it explored the relationship between extramural English and receptive and productive vocabulary size, and third, it took learners' perspectives on EE and (vocabulary) learning into account by providing a detailed analysis of their views on the importance of English in their everyday lives, their beliefs about and experiences with learning from EE, and their thoughts on the link between in- and out-of-school language learning.

To achieve these aims, a cross-sectional study using a mixed methods approach was conducted. The sequential QUAN-qual research design consists of a larger quantitative strand followed by a more in-depth qualitative exploration, which confirms, complements and enhances the quantitative data. Integration of the quantitative and qualitative strands occurred through planning, sampling, instrument development and data analysis in which quantitative and qualitative findings were drawn together in meta-inferences. In the quantitative strand, data were collected from 224 10th-grade learners of English using a detailed questionnaire (EEQ), an online language diary (EEOLD) and two vocabulary tests: V_YesNo (Meara 2015a) was used to measure written receptive vocabulary size at the level of meaning recall and Lex30 (Meara & Fitzpatrick 2000) was used to test productive vocabulary size operationalized as recall of written word form. After the application of exclusion criteria, a final sample of 201 participants could be used in the quantitative analysis. The quantitative sample also formed the basis for participant recruitment of the smaller qualitative strand, which used focus group interviews to elicit students' views on extramural English and its learning potential. In total, six focus groups

with 30 participants were conducted; these were then transcribed and analysed using qualitative content analysis. Consequently, the quantitative and qualitative data were first analysed individually before they were integrated in a final, separate stage of analysis. Using such a sequential MMR design to study extramural English helps to improve the quality of inferences in at least two ways: it allows to collect and compare quantitative and qualitative data on the same constructs and it allows to discuss the quantitative results with some of the participants who contributed to them akin to member checking in qualitative research (Creswell & Miller 2000).

The study forms part of the emerging research area of language learning beyond the classroom, which has received growing interest since the late 1990s, particularly in relation to informal language learning in out-of-school contexts. Since it is still in the early stages of development, the field is currently characterized by a variety of theoretical concepts and approaches originating in the diverse interests and backgrounds of researchers. Chapter 2 discusses the most relevant approaches by comparing them to Benson's (2011) model of language learning beyond the classroom and to extramural English (Sundqvist 2009) as the central theoretical concept used in this thesis. To give an overview of the scope of existing research on informal out-of-school language learning and to identify areas in need of further investigation, a synthesis of studies along the five dimensions of time, location, conceptual background, content focus, and research design is provided.

Chapter 3 is concerned with the second focus of this study and reviews previous research on EE and vocabulary learning. It also provides more general information on L2 vocabulary development and critically discusses the foundations of vocabulary research because the range of different approaches has important implications for vocabulary testing and the interpretation of results. The complex nature of vocabulary measurement, which stands in contrast to the apparent simplicity of vocabulary tests, is highlighted in relation to vocabulary size before summarizing the results of research in relation to EE. Findings suggest that there is a positive relationship between learners' overall engagement with extramural English and vocabulary knowledge; in addition, research focusing on vocabulary learning from specific activities like reading, listening, viewing and gaming shows that vocabulary can be acquired from all of these, although learning gains in incidental acquisition are typically small.

Chapter 4 provides information on English in Austria to contextualize the findings of the present study. After a brief introduction to the linguistic situation, a discussion of the roles of English shows that despite having no official status, English clearly is in a special position because next to (Austrian) German it is the dominant language in education, business, the linguistic landscape and the media. Data on leisure activities and media use among adolescents indicate that Austrian teenagers have access to a wide range of media and the results of the small body of existing research on EE suggest that young Austrians from primary to tertiary level use English at least to some extent for spare time activities such listening to music or watching audiovisual media. Finally, the specific research context of the present study is characterized using Benson's

model of language learning beyond the classroom before presenting the empirical study and its results in the remaining chapters.

Chapter 5 sets out the specifics of the MMR design after providing information on the methodological approach of mixed methods research. It presents and explains decisions taken in relation to sampling, instrument development and selection, as well as data collection and analysis for each of the two strands and contains detailed information on all research steps including piloting. The results of the quantitative analysis are presented in Chapter 6: data on participants' background and their perceptions of English are summarized before presenting the findings on the types and amount of extramural English. For receptive and productive vocabulary size, methodological aspects are investigated first before exploring their relationship with EE and further potential influencing factors using both bivariate and multivariate analysis techniques. In addition, findings from a new method of analysis for Lex30, which compared the response words produced in the test against participants' respective English coursebooks, are presented. Chapter 7 reports the results of the qualitative strand in relation to the main themes identified in the qualitative content analysis. First, participants' views on the significance of English in the everyday lives of young Austrians are described followed by more in-depth information on their EE practices. Learners' experiences with and evaluations of learning in extramural contexts constitute the second focus of this chapter; here, special attention is paid to their practices in relation to new vocabulary encountered in EE activities. Finally, participants' accounts of using and learning English inside and outside school provide insights into their perceptions of these two learning environments.

The quantitative and qualitative findings presented in Chapters 6 and 7 are then drawn together in a final stage of analysis. Chapter 8 first presents conceptual and methodological contributions of this thesis in relation to extramural English, the use of mixed methods, and vocabulary measurement. In a second step it discusses the integrated empirical results in relation to the three main research aims and compares the outcomes of the present study to those of previous research.

9.2 Key findings and implications

In the following, key findings are highlighted in relation to each of the five main research questions that guided the design of the mixed methods study (see section 5.1).

> **RQ 1:** *What kinds of extramural contacts do Viennese upper secondary school students report having with English?*

Viennese teenagers engage in many different activities during their leisure time and the data show that what they like doing in general, they also do a lot in English. The three most popular EE activities in which over 50% of the 201 participants engage (almost) every day are listening to music, watching online video clips and reading in social media, which are also the three most common general leisure activities. These few widespread activities stand in contrast to the overall nature of participants' EE environments, which are highly individualized and impressively varied: all 64 EE activities listed in the Extramural English Questionnaire (EEQ) are

done by at least one participant at least a few times a month. Moreover, both in the quantitative and qualitative data participants report further niche activities through which they pursue specialized interests.

The most frequent EE activities can be grouped into three broad categories: first, participants engage with English-language music through listening and singing; second, they watch 'online TV' in the form of English-language video clips, movies and series; and third, they read, and to an extent use, English in other online environments such as social media, apps or search engines. Consequently, popular EE activities among Viennese teenagers typically involve a screen, most are carried out online, and almost all mainly involve language in a receptive way. In contrast, specialized niche activities, which range from creative writing over programming and rapping to participating in English-speaking sports teams, entail more productive language use.

In relation to the types of extramural contacts, the data further indicate that there are differences according to gender. Male participants play all types of digital games significantly more often and also use different forms of game-related communication, such as in-game chats or VOIP services, more frequently. Female participants engage more frequently with music and related activities such as singing or reading lyrics, they read books or stories significantly more often and they are also more likely to write stories or a diary in English.

In response to the RQ 1, we can thus conclude that Viennese upper secondary school students

- … mainly come in contact with extramural English through music, 'online TV' in form of video clips, films and series, and other online environments such as social media, apps and search engines.
- … construct diverse EE environments in line with their personal preferences and specialized interests.
- … mostly use the receptive language skills of listening and reading for their EE activities.

RQ 2: *How much time do Viennese upper secondary school students report spending in contact with extramural English?*

The questionnaire data show that the extent of contact with extramural English is remarkable with over 96% of participants reporting at least one EE activity per day. This finding is complemented by more specific data based on the Extramural English Online Language Diary (EEOLD) which suggest that Viennese adolescents spend a substantial amount of their leisure time with English: on average, participants engage in EE activities for just over four hours a day. This finding is supported by students participating in the focus group interviews, who suggest that music and the use of smartphones, which allow 'doing EE on the move', could be factors contributing to this large amount of time. However, it also needs to be acknowledged that there is great variation in relation to time spent with extramural English as indicated by the large standard deviation of the mean, which amounts to over two and a half hours. In terms of skills, most time is spent on listening and viewing activities, which last for approximately two hours, followed by reading with just below one hour. Speaking, writing and multi-skill activities such as gaming are carried out for less than 30 minutes per day on average.

Furthermore, there is a statistically significant gender difference with regard to mean time spent with EE: male participants spend significantly more time per day with English in out-of-school contexts than female participants, while the difference according to frequency of EE activities, which is also higher for boys, is not statistically significant. Further influencing factors that show statistically significant associations with frequency of EE engagement in the bivariate analysis are socioeconomic status and the related variables of media access and the number of books available at students' homes. In addition, overall self-assessed English proficiency also shows a significant positive correlation.

The response to RQ 2 thus shows that Viennese upper secondary school students
- … vary greatly in relation to the amount of time they spend with extramural English.
- … on average spend just over four hours with EE activities per day.
- … spend significantly more time with EE if they are male.

RQ 3: *What is the relationship between extramural English and the receptive English vocabulary size of Viennese upper secondary school students?*

The Viennese 10th-grade students participating in this study have an estimated mean receptive vocabulary size of approximately 4,800 lemmas based on V_YesNo in combination with the S-shaped logistic weighting function recommended for scoring this test (Meara & Miralpeix 2017). However, the test can also be scored using a second, stricter scoring method: the *h×CJ%* score uses each individual participant's proportion of correct judgements, referring to corresponding responses on 20 items in the V_YesNo test and 20 translation items of the respective target words, to correct their raw number of hits. Hence, in contrast to other scoring formulae for Yes/No tests, the proportion of correct judgements in relation to the translation items takes evidence for participants' knowledge of the target words into account. Based on the stricter *h×CJ%* score the participants have a mean receptive vocabulary size of 3850 lemmas, which is considerably lower than the V_YesNo score. Consequently, all further analyses were carried out and reported for both sets of scores, although most results are very similar.

The relationships between receptive vocabulary size, extramural English and other potential influencing factors were explored on the basis of the combined dataset. In a first step, bivariate analysis shows a statistically significant correlation between receptive vocabulary size and frequency of engagement with EE. Although the effect is small, it persists in the multivariate analysis: in the standard multiple regression model for receptive vocabulary size, the EE median score and length of English instruction operationalized as the number of years spent learning English emerge as the two significant predictors for both V_YesNo and *h×CJ%* as outcome variables. Overall, the model, which also contains SES, media access, gender and the number of home languages as independent variables in addition to the two significant predictors, explains 21.3% of variance in participants' test scores. What is intriguing is that frequency of engagement with EE on its own explains slightly more variance than length of instruction. This finding could be taken to mean that current engagement with extramural English has a greater effect on

receptive vocabulary size than the number of years spent learning at school, although it has to be stressed that both effects are very small.

In addition to EE and length of instruction, the three SES-related variables also show statistically significant positive correlations with receptive vocabulary size with comparable, small effects in the bivariate analysis, while the number of home languages is significantly negatively related to receptive vocabulary knowledge. However, closer inspection reveals that the relationship between receptive vocabulary size and the number of home languages is mediated by SES; hence, growing up multilingually is not actually detrimental to acquiring a large English vocabulary, but correlates with a lower SES.

Concerning RQ 3 and the relationship between extramural English and receptive vocabulary size, the results show

 … a statistically significant positive relationship between the frequency of engagement with EE and the two receptive test scores in both bivariate and multivariate analyses, even though the effect is small.

 … that current frequency of engagement with EE explains slightly more variance than length of English instruction in the multiple regression model.

RQ 4: What is the relationship between extramural English and the productive English vocabulary size of Viennese upper secondary school students?

Productive vocabulary size was measured using Lex30, which does not produce a score that can be interpreted in terms of a concrete size estimate. However, a comparison with outcomes of other studies indicates that the participants scored similarly to or slightly higher than comparable samples of learners. The relationship between Lex30 scores, EE and other potential influencing factors was again explored with the help of the combined dataset. In contrast to receptive vocabulary knowledge, frequency of contact with EE does not show a statistically significant relationship with productive vocabulary size either in the bivariate or the multivariate analysis. The only statistically significant predictor in the standard multiple regression model for productive vocabulary size is length of English instruction; this variable thus predicts both receptive and productive vocabulary size, which is not unexpected. With regard to further factors, a statistically significant difference of Lex30 scores according to gender was found in the bivariate analysis. In addition, the SES summary variable as well as the number of books available at students' homes are significantly and positively correlated with Lex30 scores, but the effects are very small. However, none of these effects emerge in the presence of other predictors in the multivariate analysis.

The finding that length of instruction is the only statistically significant predictor of productive vocabulary size and the lack of a relationship with frequency of engagement with extramural English becomes plausible when considering that most popular EE activities only involve language in a receptive way. What is, however, more surprising is that there was no statistically significant difference in productive vocabulary size between students who engage in 'niche activities', which entail more productive language use, and the remaining participants either.

The quantitative data indicate that there is no relationship between EE and productive vocabulary size, and yet, two other findings suggest that some productive vocabulary knowledge is acquired extramurally. The first result is that participants in the focus group interviews are able to provide examples of lexical learning from EE. Since the students can reproduce the words given as examples of learning, they must have acquired some productive knowledge from EE. The second finding relates to the exploratory schoolbook analysis, in which response words produced on the Lex30 test were compared to learners' past and present coursebooks at school. Although 'extra-coursebook' is a rather imprecise operationalization of extramural as teachers commonly teach more than the contents of the coursebook, this comparative analysis allows a more in-depth exploration going beyond the conventional analysis of Lex30 data. It shows that 20% of the response words produced by participants were not included in their respective schoolbooks and most of these belong to the category of mid-frequency vocabulary, which includes general-purpose vocabulary met relatively frequently in authentic language input (Nation 2013). Furthermore, a categorization of the offlist types according to thematic fields shows that many relate to topics typically encountered in out-of-school leisure activities such as fantasy, historical warfare, or crime, terrorism and modern warfare. Since these are the themes explored in many films, series and games, it is very likely that at least some of the response words produced in the Lex30 test but not found in participants' schoolbooks were acquired from EE.

The contradictory findings of the statistical analyses, the qualitative interview data and the schoolbook analysis could also indicate that measurement issues mask the existence of a relationship between EE and productive vocabulary size and that Lex30 is not a sensitive enough test to investigate this connection. The findings of the present study suggest that the relation between EE and productive vocabulary is qualitatively different from and much weaker than the relation with receptive vocabulary size, but clearly further research is needed in this respect.

With regard to RQ 4 and the relationship between extramural English and productive vocabulary size, the analysis shows

- … that there is no statistically significant relationship between frequency of engagement with EE and productive vocabulary size either in the bivariate or in the multivariate analysis.
- … that participants are able to provide examples of lexical items acquired from EE in the focus group interviews, which indicates that at least some productive knowledge can be gained from engagement with EE.
- … that at least some of the words have likely been acquired from EE activities as indicated by an exploratory analysis that compared response words on the Lex30 test to participants' past and present coursebooks.

> *RQ 5: What are Viennese upper secondary school students' perceptions of EE and its potential for language learning?*

This qualitatively-oriented research question is the broadest in scope and was split into seven more specific sub-questions to guide the development of the interview guide used in the focus groups (see Chapter 5). A short synthesis therefore cannot do justice to the many insights provided by participants in the interviews and the multifaceted nature of the qualitative data, but an attempt has been made to briefly summarize the most pertinent findings on Viennese adolescents' perceptions of the significance of English for young Austrians, their view on learning from EE with a special focus on vocabulary acquisition, and their thoughts on the relationship between in- and out-of-school English practices.

Both the qualitative and quantitative data indicate that English is important for the participants: it is regarded as the most significant language in addition to their first language(s) and it is used more frequently in their leisure time than any other additional language. Indeed, knowing and using English appear to have become 'normal' for many adolescents, although the importance attributed to it in young Austrians' daily lives varies from seeing it as absolutely necessary to evaluating it as an added benefit. There are two main reasons for the significance of English: its role as the global lingua franca, or 'universal language' as one participant characterized it, which is seen as impacting international travel, worldwide communication and future jobs, and the fact that it is needed for leisure activities. Hence, EE actually contributes to the perceived importance of English among Viennese teenagers. The primary reasons given for using English for leisure activities in extramural contexts are a preference for original versions, which is in accordance with the aesthetic qualities ascribed to English, and the greater and earlier availability of content in comparison to other languages.

With regard to the question of learning, a large majority of participants in the focus groups believe that they benefit from engagement with EE in terms of language development. Many participants hold a strong and unquestioned belief that communication with and authentic input from native speakers is most helpful for learning English and that, conversely, picking up structures from people who use 'bad', incorrect language is the main obstacle of learning from EE. Further factors that are seen as contributing to language development relate to repetition, both in the sense of repeated encounters with a language structure and repeated engagement in an EE activity, a familiarization effect which facilitates comprehension over time, and the positive effect of the voluntary and enjoyable nature of informal activities on affective variables such as motivation. In contrast, exposure to incorrect grammatical structures and slang in media such as pop songs as well as wrong inferences are identified as further problems with learning from EE; overall, however, participants name more positive than negative factors in line with the majority's belief that EE has positive effects on the development of English proficiency.

Participants' accounts of vocabulary learning from EE as well as the questionnaire data indicate that guessing from context is the default strategy to discover the meaning of unknown lexical items and that additionally comparisons with other languages and consultations of online

dictionaries are used if required. Dictionaries are subject to a controversy because some students see using them as an easy and fast option to find out the exact meaning of a new word, whereas for others dictionaries are the very last resort that is only used if all other strategies have failed. In addition, participants are very clear about the fact that not all new lexical items warrant strategic behaviour. Lexical inferencing is used for many new words, but in order to receive further attention unknown words or phrases need to have one of three characteristics: they are important for understanding the content *and* difficult to infer from context, they have been met repeatedly in EE input, or they somehow attract interest, for instance through formal properties.

Lastly, the focus groups also provided insights into participants' perceptions of the relationship between English inside and outside school. As in previous research, the majority do not see a link between English teaching and extramural English, but they argue that English lessons at school formed the basis for their EE activities and any informal learning. At the same time, several students state that in their opinion engagement with EE has contributed more to their knowledge of English than lessons at school, particularly at upper secondary level.

In response to RQ 5 on students' perceptions of extramural English and its potential for language learning, the main conclusions are that

- … participants evaluate English as the most important language for young Austrians in addition to their L1(s) and that EE contributes to this perceived significance.
- … EE is beneficial to language development in most participants' view.
- … guessing from context is the default strategy for unknown words encountered in EE activities and that new lexical items need to have certain characteristics to warrant further strategic attention.
- … currently there is no link between English inside and outside school for most participants, but many students argue that lessons at school provided the foundation for extramural engagement with English.

The overall MMR research question of the present study was

> ***What is the impact of extramural English on Viennese upper secondary school students' vocabulary knowledge and development?***

Drawing together all the findings reported above and acknowledging the limitations of the context, sample and instruments of the empirical study, we can conclude that

- … frequency of engagement with EE has a measurable and statistically significant impact on receptive vocabulary size.
- … there is no measurable quantitative effect of EE on productive vocabulary size, although the qualitative data and the exploratory schoolbook analysis suggest that some productive knowledge can be acquired from EE.
- … therefore, EE activities among Viennese 10th-grade students appear to affect receptive and productive vocabulary size differently.
- … EE has a large perceived impact on English language learning in general and vocabulary learning in particular as identified by the learners.

9.3   Significance, limitations and outlook

The mixed methods study presented in this thesis is concerned with the "specific local characteristics" (Sockett 2014: 156) of informal language learning from extramural English among 15- to 16-year-old teenagers in Vienna, Austria. As such, it is valuable from an Austrian perspective because it provides much needed data on the EE activities of upper secondary school students and, together with its partial replication by Hahn (2017, 2018), it is the first study to systematically explore the relationship between engagement with EE and language gains in form of vocabulary knowledge in this context.

However, the study is informative to a much wider audience beyond the local Austrian context for at least four reasons. First, it allows to draw comparisons between this relatively new research environment and previous research in terms of EE exposure and vocabulary learning. As shown in Chapter 2, much research on extramural English has been conducted in so-called subtitling countries, in which children are exposed to large amounts of English input through subtitled television from an early age onwards, and this is especially true for studies investigating the relation between EE and vocabulary learning (De Wilde, Brysbaert & Eyckmans 2019; De Wilde & Eyckmans 2017; Hannibal Jensen 2017; Jóhannsdóttir 2017; Kuppens 2010; Olsson 2012; Olsson & Sylvén 2015; Persson & Prins 2012; Peters 2018; Peters et al. 2019; Puimège & Peters 2019; Sundqvist 2009a; Sylvén 2004/2010; Verspoor, De Bot & Van Rein 2011). Although EE research in non-subtitling countries that use dubbing or voice-over practices for foreign language TV broadcasts is growing (e.g. Grau 2009; Kusyk 2017; Kusyk & Sockett 2012; Mirmán Flores & García Jiménez 2018; Muñoz 2012; Sockett 2013; Toffoli & Sockett 2013), few of these projects focus on vocabulary and thus do not allow a comparative analysis of the effects of early television exposure. A comparison of this study's findings to previous research shows that in terms of engagement with EE the difference between subtitling countries and Austria as a dubbing country is negligible because traditional TV only plays a minor role in the media environments of young people these days: Viennese teenagers mainly watch films and series through online streaming platforms on which they can choose the language settings and, as we have seen, many opt to use English for reasons of availability or a preference for original versions.

However, the same is not true for vocabulary knowledge because in this respect early exposure seems to make more of a difference: the vocabulary sizes reported in studies from Belgium (e.g. Peters 2018) or Denmark (Henriksen 2008; Stæhr 2009) tend to be greater than those found in subtitling countries like Austria, Germany or Spain after comparable lengths of instruction (see Table 3.4 in section 3.1.2), although it has to be noted that measurement issues render such comparisons difficult and inexact. In addition, studies with young language learners in Belgium (De Wilde & Eyckmans 2017; De Wilde, Brysbaert & Eyckmans 2019; Puimège & Peters 2019) and the Netherlands (Persson & Prins 2012) have established that children can acquire a sizeable amount of English vocabulary from exposure to EE before they start formal English instruction, while the qualitative findings of this study show that many Viennese students consider English

teaching at school to be the foundation for their engagement with EE. Hence, this study adds to the existing body of research by showing that differences between subtitling countries with early exposure to English and dubbing countries like Austria do not play a role with regard to adolescent learners' current engagement with extramural English, but they do make a difference in relation to the beginning of EE activities and the trajectory of informal language learning, and they also seem to affect vocabulary size.

A second strength of this study in relation to research on EE and vocabulary knowledge lies in its focus on learner perspectives. To date, little research (e.g. Anioł 2011; Grau 2009; Ingvarsdóttir & Jóhannsdóttir 2017; Lai 2015) has collected detailed qualitative data on adolescents' views on extramural English and informal language learning and to the best of my knowledge there are no studies that explicitly discuss contextual vocabulary learning from EE with learners, although they are the most important agents in the (vocabulary) learning process. The qualitative strand of this MMR study allows new insights into learners' practices of and beliefs about vocabulary learning. The discussion mainly focused on the acquisition of meaning as this is the most prominent aspect of vocabulary learning for participants, although they mention that knowledge of written and spoken word forms can also be learned from EE input. As described in the previous section, participants' accounts indicate that only few strategies are used to discover the meaning of unknown lexical items with contextual inferencing being the default option. In addition, they appear to have specific, though often implicit, criteria for which new words warrant strategic attention. These findings are relevant to research on incidental vocabulary learning in authentic, real-life circumstances and highlight the value of taking learners' views into account.

Further innovations introduced by this study are of a methodological nature. As stated at the beginning of this chapter, using a sequential MMR design to collect both quantitative and qualitative data on the research topic allows to explore it in greater depth by comparing different perspectives and to further investigate unexpected findings. In addition to the overall research design, the study also explored new ways of scoring the data collected with the help of V_YesNo and Lex30. With regard to the receptive vocabulary test, a stricter scoring method for Yes/No data that takes evidence for word knowledge into account by comparing participants' judgements on the test to translation items was first put into practice and showed that the formula proposed by Meara and Miralpeix (2017) leads to an overestimation of receptive vocabulary size in the present sample. In relation to productive vocabulary, a statement by Fitzpatrick and Clenton (2010: 551) that Lex30 can potentially "be combined with a range of analytical measures" in addition to the standard frequency-based analysis has been taken up in this study. In the exploratory schoolbook analysis, the response words produced in the Lex30 tasks were analysed against participants' past and present English coursebooks to further investigate relations between EE and productive vocabulary learning. The focus of analysis was on words that participants were able to produce on the test, but which were not included in their respective coursebooks. Combining this additional quantitative procedure with a thematic

analysis allowed an in-depth exploration of the Lex30 data, which indicates that some productive knowledge is likely learned from engagement with EE, although the quantitative analysis suggests otherwise (see section 9.2).

Finally, this thesis also makes conceptual contributions to the emerging research area of (informal) language learning beyond the classroom. In Chapter 2 an attempt was made to disentangle the competing approaches and concepts currently used in the field and to clarify the scope and primary focus of *extramural English* (EE), *language learning beyond the classroom* (LBC), *informal digital learning of English* (IDLE) and *online informal learning of English* (OILE). A comparison of these four conceptualizations has shown that they differ in terms of their inclusiveness of contexts for language learning and use and in their perspective on the occurrence of learning. In addition, the review of existing research on EE and related concepts highlighted areas that have been neglected in terms of empirical research so far and thus points to directions for future research, as is discussed in the next section.

In summary, this thesis adds to existing research on extramural English and vocabulary learning because it explores a new learner population in a currently under-represented research context, introduces methodological innovations, provides detailed insights using several empirical perspectives, and dedicates space to learners' voices.

However, like most research, this study is subject to certain limitations, many of which relate to the object of enquiry and the naturalistic conditions in which this project was carried out. As discussed in section 8.1, EE activities can rarely be directly observed because they take place in private settings during learners' leisure time, which means that studies have to rely on self-report data. While triangulation of datasets from different instruments, as done in the present study, can help to check the accuracy of self-report data, they can never be completely verified. In addition, the fact that EE takes place in private contexts also means that researchers cannot exert any control over participants' engagement with EE, which precludes the possibility of using a control group. These two conditions present important limitations, especially for the quantitatively-oriented research strand, but they are an inevitable part of investigations into this real-life phenomenon. In response, the MMR approach used attempts to establish as comprehensive a picture as possible of 15- to 16-year-old Viennese teenagers' EE practices by examining them from several angles and collecting frequency-based and time-based quantitative data as well as qualitative reports. This in-depth exploration of a specific local context is both a strength and a weakness of this study because although it increases the quality of its inferences and adds to its ecological validity, the statistical generalizability of the results based on the present sample is limited to a population of Viennese adolescents in their mid-teens in the second half of the 2010s. It is very likely that the results are applicable to adolescents in other parts of Austria and further comparable contexts, but we cannot say so with certainty. In addition, it needs to be acknowledged that the results of this and other studies on EE will relatively soon be outdated due to the volatile nature of EE practices, which are constrained by technological conditions, media products and leisure trends that are in constant change.

A second area of limitations is concerned with the specific data collection procedures used in the present study. First, administering the language diary online emerged as problematic because for reasons of anonymity and data protection I did not have access to participants' contact details and could not regularly remind them to fill in the EEOLD. Although I stressed the importance of the language diary and hung posters with reminders in their classrooms, many participants did not regularly complete it, which resulted in a low overall response rate and limits the usefulness of the time-based data for statistical analysis. Second, it was difficult to decide on the methods of vocabulary measurement: since target words that participants have acquired through engagement with EE cannot be established in advance and are, of course, likely to vary for each individual student, tests of vocabulary size have to be used as a proxy. As discussed in Chapter 3, measuring vocabulary size is a difficult undertaking and the various options available all have advantages and disadvantages. The two tests used in the empirical study were chosen mainly for their practical advantages in relation to data collection because, as is common in school-based research, I could only gain access to students for limited amounts of time. Like other measures, V_YesNo and Lex30 have certain drawbacks that were openly discussed in section 5.3.3.3 and further explored in the quantitative data analysis. The main limitations of both tests relate to scoring issues, which were addressed through methodological innovations in this study. Furthermore, the results indicate that Lex30 may not be sensitive enough to small gains in productive vocabulary knowledge to show a measurable effect of EE, although further analyses suggest that some productive knowledge is acquired from EE activities. Evidently, further research is needed in this respect, which will be discussed below.

Lastly, there are also some limitations in relation to research design and especially with regard to quantitative data analysis because the low number of participants in some classes and schools rendered the use of more powerful statistical analysis tools such as mixed models impossible. While participant numbers in existent groups are beyond researchers' control because of the voluntary nature of participation, anticipating greater drop-out rates in relation to exclusion criteria and reliability thresholds and thus collecting more data could have helped with these issues. Moreover, the number of variables included in this study was necessarily limited and further explanatory factors could have been taken into account; in particular, more detailed data on participants' English instruction at school and the inclusion of learner factors such as agency, motivation, learning style or vocabulary learning aptitude would have been desirable but were not possible for practical reasons. At the same time, the reliability of some questionnaire scales such as attitudes towards English or use of VLS was low, which restricted their use to descriptive analysis and points to the conclusion that it may be better to include detailed measurement of a few, select variables in quantitative research. In the qualitative strand, it would have been interesting to include the teachers' perspective on the relation between EE and language teaching at school, but such an extension was not possible for reasons of scope.

Many of the limitations set out above entail calls for further investigation and since research on extramural English and informal (vocabulary) learning has developed into a thriving area of enquiry, it is to be hoped that these calls will soon be answered. Indeed, much work remains to be done in this emerging research field, as was pointed out at the end of the literature review on EE and lang/uage development in Chapter 2. As a preliminary step, the field would benefit from theoretical consolidation and further development of models such as Benson's (2011) framework because a clearer conceptual foundation would enhance the comparability of different studies. As the detailed analysis of prominent conceptual approaches undertaken in this thesis showed, there are important differences between the concepts currently used with regard to their scope and their perspectives on language learning and use. In a second step, a look towards neighbouring research strands and more long-standing discourses on informal learning in other subjects, such as those briefly described at the beginning of Chapter 2, could also contribute to further theoretical development and usefully inform the emerging research field of extramural English and informal language learning.

Further explorations of extramural practices in diverse contexts and among diverse populations are needed and, perhaps most importantly, research on target languages other than English (see also Sockett 2014). The broad overview given in Chapter 2 indicates that contexts outside Europe and Asia are in danger of being neglected and there are few studies focusing on languages other than English. Although the special position of English clearly contributes to extramural engagement because resources are widely available, that is not to say that extramural activities in other languages are not possible, as indicated by the little existing research (e.g. Bengtsson 2014; Kalaja et al. 2011; Nightingale 2016). By taking a multilingual viewpoint, comparative studies of English and other languages could shed light on similarities and differences of the learning mechanisms at play. In addition, it would be interesting to repeat previous studies after some time to assess the impact of new technologies or changing fashions on extramural (English) practices.

Regarding extramural English, further research is needed with younger participants to learn more about how, when and why they begin using English for informal activities and which factors play a role for initial EE activities in different contexts. Studies on young language learners can also help to solve the 'chicken or egg' dilemma in relation to EE and language proficiency because the development of EE practices and language proficiency could be tracked over time in longitudinal studies and could thus provide information on whether learners who engage more often with EE are more proficient, or whether more proficient or more talented learners engage more frequently with EE. Besides young learners, it would be interesting to explore the English leisure practices of adults outside formal education, although such research presents difficulties in relation to access. In addition to different groups of learners, in-depth explorations of individuals' EE practices in case studies could provide highly informative data on the specifics of language use, learners' attention to language during EE activities, and/or the characteristics of language learning in such activities. Moreover, in such case studies EE

practices could potentially be observed directly in addition to collecting different types of self-report data.[208]

Additional research should also be conducted on the effects of extramural activities on the development of language skills and different areas of linguistic knowledge. The present study focused on vocabulary knowledge and in this respect, further investigation of whether and how EE affects receptive and productive vocabulary knowledge differently is needed. Two methodological innovations introduced in this study, the *h×CJ%* score for V_YesNo and the schoolbook analysis of the Lex30 data, could also be explored further; especially, alternative methods of analysing the Lex30 data could be an interesting direction of research worth pursuing. In addition, qualitative case studies could also provide more detailed insights into vocabulary acquisition from EE, as suggested above. Schmitt (2019: 267) also suggests examining "the precise nature of exposure" in extramural contexts because

> [h]aving finer-grained detail about the nature of extramural exposure, and studying how this directly leads to L2 acquisition, would allow more concrete suggestions about how to promote the most effective use of extramural exposure in a range of contexts, and how to best integrate it with classroom instruction (Schmitt 2019: 267).

In addition, he recommends focusing on digital games as a prominent EE activity and to analyse the lexis they contain as a second research task in relation to EE. I would suggest that studies could also explore vocabulary learning and use in online writing in blogs, forums, fan fiction or even social media because, as shown in section 3.3.2.4, at present there are no studies on vocabulary acquisition in such contexts.

However, to me, the most important next step in research on extramural English concerns its pedagogical implications. Although EE does not posit learning as a given, the ultimate goal of much, if not most, EE research is to support language development across different in- and out-of-school contexts. The results of the empirical study presented in this thesis as well as those of previous research in other contexts show that learning through leisure is a reality and that it is high time to acknowledge the value of language learning outside formal educational environments. While we certainly need more data on EE practices in different contexts, age groups and school types and more detailed information on the views of the learners, it is also time to begin implementing changes based on existent research. From the research reviewed and presented in this thesis, it should have become clear that the context in which English learning and teaching nowadays take place in Austria and many other parts of the world is very different from two decades ago when YouTube, Instagram and Netflix, to name but a few favourite platforms, were not yet available. Clearly, such changes in the teaching and learning context present challenges for teachers and schools; however, in Austria and elsewhere approaches to ELT have not yet been systematically adapted to take account of the changed context, changed aims and expectations, and changed roles of both teachers and learners.

---

[208] While the focus of these suggestions is on EE, the latter two proposals are similar to tasks 1 and 2 in Reinders' and Benson's (2017) research agenda for language learning beyond the classroom. Many of their insightful suggestions in relation to this wider area of enquiry could also inform EE-focused research.

Although a detailed discussion of such issues has not been possible in this thesis for reasons of scope, first thoughts on what such changes could look like are sketched below.

The key findings of this study clearly have implications for ELT because they show that almost all participants engage with EE on a frequent basis and that on average they spend much more time with English outside school than inside school. Yet, the results also reveal that Viennese teenagers engage in a wide range of EE activities and that there is great variation in terms of frequency of activities, amount of time spent with EE, and type of activities. This finding suggests that there is no one-size-fits-all approach to taking EE activities more into account in ELT (see also Sundqvist & Sylvén 2016). The qualitative results indicate that students value informal language learning and that many of them believe that by now it contributes more to their language competence that lessons at school. However, at the same time, learners' lay theories laid out in the focus group interviews indicate that they are not always aware of their language and language learning needs and therefore it is the teachers' job to help them recognize and address them.

In doing so, a first step for teachers would be to acknowledge the existence of lifewide learning (Banks et al. 2007) and to support it in parallel to lifelong learning to legitimize informal language acquisition. In line with most initial proposal for linking in- and out-of-class learning (Caspari 2015; Nunan & Richards 2015a; Richards 2015; Sockett 2014; Sundqvist & Sylvén 2016), I would argue that classroom practice should not be modelled on students' language use and learning outside school, but endeavour, on the one hand, to facilitate and enhance such learning and, on the other hand, to usefully complement it by focusing on those skills and types of knowledge that students are unlikely to acquire in extramural contexts. Hence, rather than being *EE-inclusive*, English teaching in the 21st century should ideally be *EE-sensitive*. At the moment, this area is both under-theorized and under-researched because we lack concrete examples of how to put suggestions for EE-sensitive teaching into practice; hence, further theoretical development and school-based research is urgently needed and constitutes the most pressing research task in my view. Perhaps, some of the results and ideas presented in this thesis can act as a springboard for such research and teaching projects.

This study is but a small step toward understanding the new, hybrid learning context that constitutes extramural English, but it has shown that even in a non-subtitling country that traditionally considers English to be a foreign language, teenagers now have unprecedented amounts of contact with English and that their engagement with EE has measurable effects on at least their receptive vocabulary knowledge. It is to be hoped that further studies will explore this phenomenon in all its aspects in additional contexts because extramural English and informal language learning beyond the classroom present an exciting new area of research.

References

Abram, James & Williams, Steve. 2016. *English in context 6*: *Student's book*, 2nd edn. Linz: Veritas.

Abram, James & Williams, Steve. 2017. *English in context 5*: *Student's book*, 4th edn. Linz: Veritas.

Aitchison, Jean. 2003. *Words in the mind*: *An introduction to the mental lexicon*, 3rd edn. Malden, MA: Blackwell.

Albrechtsen, Dorte; Haastrup, Kirsten & Henriksen, Birgit. 2008. Lexical knowledge, lexical inferencing and writing. In Dorte Albrechtsen, Kirsten Haastrup & Birgit Henriksen (eds.), *Vocabulary and writing in a first and second language*. London: Palgrave Macmillan, 160–194.

Alderson, J. C. 2005. *Diagnosing foreign language proficiency*: *The interface between learning and assessment*. London: Continuum.

Alejo González, Rafael & Piquer Píriz, Ana. 2016. Measuring the productive vocabulary of secondary school CLIL students: Is Lex30 a valid test for low-level school learners? *VIAL (Vigo International Journal of Applied Linguistics)* 31, 31–53.

Al-Hoorie, Ali H. 2017. Sixty years of language motivation research: Looking back and looking forward. *SAGE Open* 7(1), 1-11.

Alm, Antonie. 2015. Facebook for informal language learning: Perspectives from tertiary language students. *The EUROCALL Review* 23(2), 3–18.

American Psychological Association (APA). n.d. *Education and socioeconomic status*: *Factsheet.* Washington, DC. https://www.apa.org/pi/ses/resources/publications/factsheet-education.pdf (20 February, 2019).

Anderson, Richard C. & Freebody, Peter. 1981. Vocabulary knowledge. In John T. Guthrie (ed.), *Comprehension and teaching*: *Research reviews*. Newark: International Reading Association, 77–117.

Anderson, Richard C. & Freebody, Peter. 1983. Reading comprehension and the assessment and acquisition of word knowledge. In Barbara A. Hutson (ed.), *Advances in reading/language research*: *A research annual*. Greenwich, CT: JAI Press, 231–256.

Anioł, Magdalena. 2011. New media and new literacies: Mapping extracurricular English language competences of Polish and Norwegian adolescents. In Maria Kaczmarek (ed.), *Health and well-being in adolescence*: *Part two. Media*. Poznań: Bogucki Wydawnictwo Naukowe, 101–124.

Anthony, Laurence. 2013. *AntWordProfiler* (1.4.0). Tokyo: Waseda University. www.laurenceanthony.net/software/antwordprofiler/.

Anthony, Laurence & Nation, I.S.P. 2017. *Picture vocabulary size test (PVST, Version 1.2.0)*. Tokyo: Waseda University. http://www.laurenceanthony.net/software/pvst (8 August, 2018).

Archan, Sabine. 2006. *Fremdsprachenbedarf und -kompetenzen*: *Unternehmensbefragung zu Ausbildungsqualität und Weiterbildungsbedarf* [Foreign language requirements and competences: Business survey on training quality and needs for continuous development]. Wien: Ibw.

Arnaud, Pierre J. L.; Béjoint, Henri & Thoiron, Philippe. 1985. A quoi sert le programme lexical? [What is the use of lexical programmes?]. *Les Langue Modernes* 79(3/4), 72–85.

Arndt, Henriette L. & Woore, Robert. 2018. Vocabulary learning from watching YouTube videos and reading blog posts. *Language Learning & Technology* 22(3), 124–142.

References

Auger, Denis. 2016. The diverse meanings of leisure. *Loisir et Société / Society and Leisure* 39(2), 173–176.

Augustín Llach, M. P. & Terrazas Gallego, Melania. 2009. Examining the relationship between receptive vocabulary size and written skills of primary school learners. *Atlantis: Journal of the Spanish Association of Anglo-American Studies* 31(1), 129–147.

Austrian Agency for International Cooperation in Education and Research (OeAD). 2014. *The Austrian education system.* Vienna. http://www.bildungssystem.at/ (14 May, 2018).

Bachman, Lyle F. & Palmer, Adrian S. 1996. *Language testing in practice*: *Designing and developing useful language tests*. Oxford: Oxford University Press.

Bailly, Sophie. 2011. Teenagers learning languages out of school: What, why and how do they learn? How can school help them? In Phil Benson & Hayo Reinders (eds.), *Beyond the language classroom.* Houndsmill, Basingstoke: Palgrave Macmillan, 119–131.

Banks, James A.; Au, Kathryn H.; Ball, Arnetha F.; Bell, Philip; Gordon, Edmund W.; Gutiérrez, Kris D.; Brice Heath, Shirley; Lee, Carol D.; Lee, Yuhshi; Mahiri, Jabari; Nasir, Na'ilah S.; Valdés, Guadalupe & Zhou, Min. 2007. *Learning in and out of school in diverse environments*: *Life-long, life-wide, life-deep.* Seattle: The Learning in Informal and Formal Environments Center (The LIFE Center) & Center for Multicultural Education. http://www.life-slc.org/ (27 June, 2019).

Barbee, Matthew. 2013. Extracurricular L2 input in a Japanese EFL context: Exposure, attitudes, and motivation. *Second Language Studies* 32(1), 1–58.

Barclay, Sam. 2017. The effect of word class and word length on the decay of lexical knowledge (Paper presented at 40th annual conference of the American Association for Applied Linguistics (AAAL 2017). Portland, OR, 20 March 2017).

Barclay, Sam. 2018. Understanding the role of systematic decay as part of vocabulary acquisition (Paper presented at 41st annual conference of the American Association for Applied Linguistics (AAAL 2018). Chicago, IL, 26 March 2018).

Barcroft, Joe. 2004. Effects of sentence writing in second language lexical acquisition. *Second Language Research* 20(4), 303–334.

Barcroft, Joe. 2015. *Lexical input processing and vocabulary learning*. Amsterdam: John Benjamins Publishing Company.

Barcroft, Joe; Sunderman, Gretchen & Schmitt, Norbert. 2011. Lexis. In James Simpson (ed.), *The Routledge handbook of applied linguistics*, 1st edn. Milton Park, Abingdon: Routledge, 571–583.

Bardovi-Harlig, Kathleen & Stringer, David. 2010. Variables in second language attrition: Advancing the state of the art. *Studies in Second Language Acquisition* 32(1), 1–45.

Barrow, J.; Nakanishi, Y. & Ishino, H. 1999. Assessing Japanese college students' vocabulary knowledge with a self-checking familiarity survey. *System* 27(2), 223–247.

Bates, Douglas; Maechler, Martin; Bolker, Ben & Walker, Steve. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1), 1–48.

Bauer, Laurie & Nation, I.S.P. 1993. Word families. *International Journal of Lexicography* 6(4), 253–279.

Beeckmans, Renaud; Eyckmans, June; Janssens, Vera; Dufrane, Michel & Van de Velde, Hans. 2001. Examining the yes/no vocabulary test: Some methodological issues in theory and practice. *Language Testing* 18(3), 235–274.

Beglar, David. 2010. A rasch-based validation of the Vocabulary Size Test. *Language Testing* 27(1), 101–118.

Beglar, David & Hunt, Alan. 1999. Revising and validating the 2000 word level and university word level vocabulary tests. *Language Testing* 16(2), 131–162.

References

Beglar, David & Nation, I.S.P. 2014. Assessing vocabulary. In Anthony J. Kunnan (ed.), *The companion to language assessment*: *Volume I. Abilities, contexts, and learners*. Malden, MA: Wiley-Blackwell, 172–184.

Bell, Philip; Lewenstein, Bruce; Shouse, Andrew W. & Feder, Michael A. 2009. *Learning science in informal environments*: *People, places, and pursuits*. Washington, D.C.: National Academies Press.

Bengeleil, Nazmia & Paribakht, T. 2004. L2 reading proficiency and lexical inferencing by university EFL learners. *Canadian Modern Language Review* 61(2), 225–250.

Bengtsson, Andreas. 2014. *Watching video or studying? An investigation of the extramural activities and Japanese language proficiency of foreign language learners of Japanese.* Stockholm: Stockholm University, MA thesis.

Benson, Phil. 2001. *Teaching and researching autonomy in language learning*. Harlow: Pearson.

Benson, Phil. 2011. Language learning and teaching beyond the classroom: An introduction to the field. In Phil Benson & Hayo Reinders (eds.), *Beyond the language classroom*. Houndsmill, Basingstoke: Palgrave Macmillan, 6–16.

Benson, Phil. 2015. Commenting to learn: Evidence of language and intercultural learning in comments on YouTube videos. *Language Learning and Technology* 19(3), 88–105.

Benson, Phil & Reinders, Hayo (eds.). 2011a. *Beyond the language classroom*. Houndsmill, Basingstoke: Palgrave Macmillan.

Benson, Phil & Reinders, Hayo. 2011b. Introduction. In Phil Benson & Hayo Reinders (eds.), *Beyond the language classroom*. Houndsmill, Basingstoke: Palgrave Macmillan, 1–6.

Berns, Margie; De Bot, Kees & Hasebrink, Uwe (eds.). 2007. *In the presence of English*: *Media and European youth*. New York: Springer.

Bertram, Raymond; Laine, Matti & Virkkala, Minna M. 2000. The role of derivational morphology in vocabulary acquisition: Get by with a little help from my morpheme friends. *Scandinavian Journal of Psychology* 41(4), 287–296.

Bilican, Recep & Yeşilbursa, Aysegul Amanda. 2015. Self-regulated capacity for vocabulary learning in Turkish high school students: An experimental study. *Procedia - Social and Behavioral Sciences* 197, 83–87.

Bisson, Marie-Josée; Van Heuven, Walter J. B.; Conklin, Kathy & Tunney, Richard J. 2013. Incidental acquisition of foreign language vocabulary through brief multi-modal exposure. *PLoS ONE* 8(4), 1–7.

Black, Rebecca W. 2005. Access and affiliation: The literacy and composition practices of English-language learners in an online fanfiction community. *Journal of Adolescent & Adult Literacy* 49(2), 118–128.

Black, Rebecca W. 2008. *Adolescents and online fan fiction*. New York, Oxford: Peter Lang.

Blanke, Karen & Cornelißen, Waltraud. 2005. German adolescents' time use from 1991 to 2001: Is gender symmetry in sight? *Loisir et Société / Society and Leisure* 28(2), 511–530.

Blell, Gabriele. 2015. Lernorte und Fremdsprachenlehren und -lernen [Learning spaces and foreign language teaching and learning]. In Eva Burwitz-Melzer, Frank G. Königs & Claudia Riemer (eds.), *Lernen an allen Orten? Die Rolle der Lernorte beim Lehren und Lernen.* [Learning in all places? The role of learning spaces in teaching and learning of languages]. Tübingen: Narr Verlag, 9–18.

Bonnet, Gérard. 2004. *The assessment of pupils' skills in English in eight European countries*. A European project. http://www.eva.dk/projekter/2002/evaluering-af-faget-engelsk-i-grundskolen/projektprodukter/assessmentofenglish.pdf (27 August, 2014).

Brese, Falk & Mirazchiyski, Plamen. 2013. *Measuring students' family background in large-scale international education studies* (IERI Monograph Series Special Issue 2). Hamburg: IERInstitute.

References

Brevik, Lisbeth M. 2016. The gaming outliers: Does out-of-school gaming improve boys' reading skills in English as a second language? In Eyvind Elstad (ed.), *Educational technology and polycontextual bridging*. Rotterdam: Sense Publishers, 39–61.

Brown, Ronan; Waring, Rob & Donkaewbua, Sangrawee. 2008. Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language* 20(2), 136–163.

Brunfaut, Tineke & Schmitt, Norbert. 2018. Vocabulary & research (Paper presented at EALTA Winter School 2018. Obergurgl, 13 February 2018).

Brysbaert, Marc; Stevens, Michaël; Mandera, Paweł & Keuleers, Emmanuel. 2016. How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in psychology* 7, 1–11.

Buchholz, Barbara. 2007. *Facts & Figures im Grundschul-Englisch*: *Eine Untersuchung der verbindlichen Übung Lebende Fremdsprache an österreichischen Volksschulen* [Facts & Figures on primary school English: A study of the compulsory exercise modern foreign language at Austrian elementary schools]. Vienna: LIT Verlag.

Bundesgesetz über den Österreichischen Rundfunk: (ORF-Gesetz) [Federal Act on the Austrian Broadcasting Corporation (ORF Act)]. In *BGBl. Nr. 379/1984 as amended by BGBl. I Nr. 115/2017.*

Bundesgesetz über die Förderung der außerschulischen Jugenderziehung und Jugendarbeit: (Bundes-Jugendförderungsgesetz) [Federal Law on the promotion of out-of-school youth education and youth employment (Federal youth promotion law)]. In *BGBl. I Nr. 126/2000 as amended by BGBl. I Nr. 136/2001.*

Bundesgesetz über die Vertretung der Anliegen der Jugend: (Bundes-Jugendvertretungsgesetz) [Federal Law on the representation of youth concerns (Federal youth representation law)]. In *BGBl. I Nr. 127/2000 as amended by BGBl. I Nr. 136/2001.*

Bundeskanzleramt Österreich [Federal Chancellery]. 2018. *Volksgruppen* [Ethnic minorities]. Vienna. https://www.bundeskanzleramt.gv.at/volksgruppen (9 May, 2018).

Bundesministerium für Bildung [Federal Ministry of Education]. 2014. *Lehrplan der Handelsschule* [Curriculum of the vocational business middle school]. In BGBl. II Nr. 209/2014 (Anlage B1). Vienna. www.ris.bka.gv.at/.

Bundesministerium für Bildung [Federal Ministry of Education]. 2016/17. *Informationsblätter zum Thema Migration und Schule, Nr. 5/2016-17*: *Der muttersprachliche Unterricht in Österreich: Statistische Auswertung für das Schuljahr 2015/16 by Ines Garnitschnig* [Informationsheets on the topic of migration and school. Mother tongue teaching in Austria: Statistical analysis for the school year 2015/16]. Vienna. http://www.schule-mehrsprachig.at/index.php?id=84 (14 April, 2018).

Bundesministerium für Bildung & Bundesministerium für Wissenschaft, Forschung und Wirtschaft [Federal Ministry for Education; Federal Ministry of Science, Research and Economy]. 2017. *Austrian educational system.* Vienna. bildung.gv.at (9 May, 2018).

Bundesministerium für Bildung und Frauen [Federal Ministry of Education and Women]. 2015/16. *Informationsblätter zum Thema Migration und Schule, Nr. 1/2015-16*: *Gesetzliche Grundlagen schulischer Maßnahmen für SchülerInnen mit anderen Erstsprachen als Deutsch: Gesetze und Verordnungen* [Informationsheets on the topic of migration and school. Legal foundations for educational measures for students with an L1 other than German: Laws and regulations]. Vienna. http://www.schule-mehrsprachig.at/index.php?id=84 (14 April, 2018).

Bundesministerium für Bildung, Wissenschaft und Forschung Austrian Federal Ministry of Education, Science and Research. n.d.-a. *Schulen Online* [Schools online]. Vienna. https://www.schulen-online.at/sol/index.jsf (9 February, 2016).

References

Bundesministerium für Bildung, Wissenschaft und Forschung [Federal Ministry of Education, Science and Research]. n.d.-b. *Standardisierte Reife- und Diplomprüfung*: *Rechtliche Grundlagen* [Standardized school leaving and diplom examination: Legal foundations]. Vienna. https://www.srdp.at/rechtliches/ (14 May, 2018).

Bundesministerium für Bildung, Wissenschaft und Forschung [Federal Ministry of Education, Science and Research]. 2019. *Detailergebnisse der standardisierten Reife-/Reife und Diplomprüfung*: *Haupttermin 2018/19* [Detailed results of the standardized school-leaving examination: Main exam date 2018/19]. Vienna. https://bmbwf.gv.at/fileadmin/user_upload/Pressemeldungen/PK_Unterlage__Ergebnisse_Zentralmatura_2019.pdf (16 August, 2019).

Bundesministerium für Bildung, Wissenschaft und Kultur [Federal Ministry of Education, Science and Culture]. 2004. *Lehrplan der Allgemeinbildende Höheren Schule: Oberstufe* [Curriculum of the academic secondary school: Upper secondary level]. In BGBl. II Nr. 277/2004. Vienna. https://www.ris.bka.gv.at/ (14 May, 2018).

Bundesministerium für Familien und Jugend [Federal Ministry for Family and Youth]. 2016. *7. Bericht zur Lage der Jugend in Österreich*: *Teil A. Wissen um junge Menschen in Österreich* [7th report on the situation of youth in Austria: Part A. Knowledge of young people in Austria]. Vienna. https://www.frauen-familien-jugend.bka.gv.at/jugend/jugendforschung/jugendbericht/siebter-bericht-zur-lage-der-jugend-in-oesterreich-2016.html (6 June, 2018).

Bundesministerium für Inneres [Federal Ministry of the Interior]. 2017. *Vorläufige Asylstatistik*: *Dezember 2017* [Preliminary asylum statistics: December 2017]. Vienna. https://www.bmi.gv.at/301/Statistiken/files/2017/Asylstatistik_Dezember2017.pdf (9 May, 2018).

Bundesministerium für Unterricht und kulturelle Angelegenheiten [Federal Ministry for Education, Science and Culture]. 2000. *Lehrplan der Allgemeinbildenden Höheren Schulen: Unterstufe* [Curriculum of the the academic secondary schools: Lower secondary level]. In BGBl. II Nr. 133/2000. Vienna. https://www.ris.bka.gv.at/ (14 May, 2018).

Bundesministerium für Unterricht, Kunst und Kultur [Federal Ministry of Education, Arts and Culture]. 2012. *Lehrplan der neuen Mittelschule* [Curriculum of the new middle school]. In BGBl. II Nr. 185/2012. Vienna. https://www.ris.bka.gv.at (14 May, 2018).

Bundesministerium für Unterricht, Kunst und Kultur; Bundesministerium für Wissenschaft und Forschung & Österreichisches Sprachen-Kompetenz-Zentrum [Federal Ministry for Education, the Arts and Culture; Federal Ministry for Science and Research & Austrian Centre for Language]. 2007. *Language education policy profile. Country report Austria*: *Language and language education policies in Austria.* Vienna. https://ec.europa.eu/migrant-integration/librarydoc/language-and-language-education-policies-in-austria (7 May, 2018).

Bundes-Verfassungsgesetz: (B-VG) [Federal Constitutional Law]. In *BGBl. Nr. 1/1930 as amended by BGBl. I Nr. 22/2018*.

Bunting, Leona & Lindström, Berner. 2013. Framing English learning at the intersection of school and out-of-school practices. *Journal of International Scientific Publications: Language, Individual & Society* 7(1), 205–221.

Burwitz-Melzer, Eva; Königs, Frank G. & Riemer, Claudia (eds.). 2015. *Lernen an allen Orten? Die Rolle der Lernorte beim Lehren und Lernen* [Learning in all places? The role of learning spaces in teaching and learning of languages]. Tübingen: Narr Verlag.

Busch, Brigitta & Peissl, Helmut. 2003. Sprachvielfalt im Wohnzimmer: Sprachenpolitik und Medien [Linguistic diversity in the living room: Language policy and media]. In Brigitta Busch & Rudolf De Cillia (eds.), *Sprachenpolitik in Österreich*: *Eine Bestandsaufnahme.* [Language policy in Austria: An inventory]. Frankfurt am Main: Peter Lang, 180–195.

References

Buttaroni, Susanna. 2013. Frühe Mehrsprachigkeit in der Elementarbildung [Early multilingualism in elementary education]. In Rudolf De Cillia & Eva Vetter (eds.), *Sprachenpolitik in Österreich*: *Bestandsaufnahme 2011.* [Language policy in Austria: Review 2011]. Frankfurt am Main: Peter Lang, 94–126.

Bytheway, Julie. 2015. A taxonomy of vocabulary learning strategies used in massively multiplayer online role-playing games. *CALICO Journal* 32(3), 508–527.

Cabot, Michel. 2016. In or out of school?: Meaningful output with digital and non-digital artefacts within personal English learning ecologies. *Nordic Journal of Digital Literacy* 11(3), 165–184.

Campbell, Julie; Aragon, Cecilia; Davis, Katie; Evans, Sarah; Evans, Abigail & Randall, David P. 2016. Thousands of positive reviews: Distributed mentoring in online fan communities. In Darren Gergle, Meredith R. Morris, Pernille Bjørn & Joseph Konstan (eds.), *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing (CSCW '16.* New York, NY: ACM Press, 689–702.

Canty, Angelo & Ripley, Brian. 2017. *boot: Bootstrap R (S-Plus) functions.* (R package version 1.3-20).

Capel, Annette. 2010. A1–B2 vocabulary: Insights and issues arising from the English profile wordlists project. *English Profile Journal* 1(1), 1–11.

Capel, Annette. 2012. Completing the English vocabulary profile: C1 and C2 vocabulary. *English Profile Journal* 3(1), 1–14.

Caro, Daniel H. & Cortés, Diego. 2012. Measuring family socioeconomic status: An illustration using data from PIRLS 2006. In Matthias Von Davier & Dirk Hastedt (eds.), *Issues and methodologies in large-scale assessments* (IERI Monograph Series Volume 5). Hamburg: IERInstitute, 9–33.

Carroll, John B. & Sapon, Stanley M. 1959. *Modern language aptitude test.* New York: Psychological Corporation.

Caspari, Daniela. 2015. Schulisches und außerschulisches Lernen verbinden: Eine (alt-)bekannte Forderung als aktuelle Herausforderung [Linking in- and out-of-school learning. A (well-)known demand as current challenge]. In Eva Burwitz-Melzer, Frank G. Königs & Claudia Riemer (eds.), *Lernen an allen Orten? Die Rolle der Lernorte beim Lehren und Lernen.* [Learning in all places? The role of learning spaces in teaching and learning of languages]. Tübingen: Narr Verlag, 29–37.

CEDEFOP. 2011. *Learning while working*: *Success stories on workplace learning in Europe.* Luxembourg: Publications Office of the European Union. http://www.cedefop.europa.eu/de/publications-and-resources/publications/3060 (16 July, 2018).

CEDEFOP. 2015. *European guidelines for validating non-formal and informal learning*: *Cedefop reference series; No 104.* Luxembourg: Publications Office of the European Union. http://www.cedefop.europa.eu/files/3073_en.pdf (16 July, 2018).

Chan, Victoria; Spratt, Mary & Humphreys, Gillian. 2002. Autonomous language learning: Hong Kong tertiary students' attitudes and behaviours. *Evaluation & Research in Education* 16(1), 1–18.

Chapelle, Carol A. 1998. Construct definition and validity enquiry in SLA research. In Lyle F. Bachman & Andrew D. Cohen (eds.), *Interfaces between second language acquisition and language testing research.* Cambridge: Cambridge University Press, 32–70.

Chen, Chuntien & Truscott, John. 2010. The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics* 31(5), 693–713.

Chen, Hao-Jan H. & Yang, Ting-Yu C. 2013. The impact of adventure video games on foreign language learning and the perceptions of learners. *Interactive Learning Environments* 21(2), 129–141.

References

Chen, Meng-Hua; Tseng, Wen-Ta & Hsiao, Tsung-Yuan. 2018. The effectiveness of digital game-based vocabulary learning: A framework-based view of meta-analysis. *British Journal of Educational Technology* 48(1), 69–77.

Cheng, Junyu & Matthews, Joshua. 2018. The relationship between three measures of L2 vocabulary knowledge and L2 listening and reading. *Language Testing* 35(1), 3–25.

Chik, Alice. 2014. Digital gaming and language learning: Autonomy and community. *Language Learning & Technology* 18(2), 85–100.

Chiu, Yi-Hui. 2013. Computer-assisted second language vocabulary instruction: A meta-analysis. *British Journal of Educational Technology* 44(2), E52-E56.

Choi, Sungmook; Kim, Jingu & Ryu, Kwangmin. 2014. Effects of context on implicit and explicit lexical knowledge: An event-related potential study. *Neuropsychologia* 63, 226–234.

Chusanachoti, Ruedeerath. 2009. *EFL learning through language activities outside the classroom*: *A case study of English education students in Thailand.* East Lansing, MI: Michigan State University, PhD thesis.

Cohen, Andrew D. & Macaro, Ernesto (eds.). 2007. *Language learner strategies*: *Thirty years of research and practice*. Oxford: Oxford University Press.

Cohen, Jacob; Cohen, Patricia; West, Stephen G. & Aiken, Leona S. 2003. *Applied multiple regression/correlation analysis for the behavioral sciences*, 3rd edn. Mahwah, NJ: Lawrence Erlbaum Associates.

Cohen, Louis; Manion, Lawrence & Morrison, Keith. 2011. *Research methods in education*, 7th edn. London: Routledge.

Cole, Jason & Vanderplank, Robert. 2016. Comparing autonomous and class-based learners in Brazil: Evidence for the present-day advantages of informal, out-of-class learning. *System* 61, 31–42.

Colley, Helen; Hodkinson, Phil & Malcom, Janice. 2003. *Informality and formality in learning*: *A report for the Learning and Skills Research Centre*. Leeds: Learning and Skills Research Centre.

Coşkun, Abdullah & Mutlu, Hakan T. 2017. Investigating high school students' use of extramural English: A scale development study. *Journal of the Human and Social Science Researches* 6(1), 571–590.

Council of Europe. 2001. *Common European Framework of Reference for languages*: *Learning, teaching, assessment*. http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf (12 December, 2010).

Council of Europe. 2002. *Presidency conclusions*: *Barcelona European council, 15 and 16 March 2002*. http://aei.pitt.edu/43345/ (15 May, 2018).

Coxhead, Averil; Nation, I.S.P. & Sim, Dalice. 2015. Measuring the vocabulary size of native speakers of English in New Zealand secondary schools. *New Zealand Journal of Educational Studies* 50(1), 121–135.

Coxhead, Averil J. & Walls, R. 2012. TED Talks, vocabulary, and listening for EAP. *The TESOLANZ Journal* 20, 55–58.

Craik, Fergus I.M. & Lockhart, Robert S. 1972. Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior* 11(6), 671–684.

Creamer, Elizabeth G. 2018. *An introduction to fully integrated mixed methods research.* Thousand Oaks, CA: SAGE.

Creswell, John W. & Miller, Dana. 2000. Determining validity in qualitative inquiry. *Theory Into Practice* 39(3), 124–130.

Creswell, John W. & Plano Clark, Vicki L. 2007. *Designing and conducting mixed methods research.* Thousand Oaks, CA: SAGE.

References

Creswell, John W. & Plano Clark, Vicki L. 2011. *Designing and conducting mixed methods research*, 2nd edn. Thousand Oaks, CA: SAGE.

Creswell, John W.; Plano Clark, Vicki L.; Gutmann, Michelle L. & Hanson, William E. 2003. Advanced mixed methods research designs. In Abbas Tashakkori & Charles Teddlie (eds.), *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: SAGE, 209–240.

Cronbach, Lee J. 1942. An analysis of technique for diagnostic vocabulary testing. *The Journal of Educational Research* 36(3), 206–217.

Crossley, Scott A.; Salsbury, Tom; McNamara, Danielle S. & Jarvis, Scott. 2011. Predicting lexical proficiency in language learner texts using computational indices. *Language Testing* 28(4), 561–580.

Crowley, Kevin; Pierroux, Palmyre & Knutson, Karen. 2014. Informal learning in museums. In Keith R. Sawyer (ed.), *The Cambridge handbook of the learning sciences*, 2nd edn. Cambridge: Cambridge University Press, 461–478.

Crystal, David. 2003. *English as a global language*, 2nd edn. Cambridge: Cambridge University Press.

Cumming, Geoff. 2012. *Understanding the new statistics*: *Effect sizes, confidence intervals, and meta-analysis*. New York: Routledge.

Cumming, Geoff. 2014. The new statistics: Why and how. *Psychological Science* 25(1), 7–29.

Daller, Helmut; Milton, James & Treffers-Daller, Jeanine. 2007. Editors' introduction: Convention, terminology and an overview of the book. In Helmut Daller, James Milton & Jeanine Treffers-Daller (eds.), *Modelling and assessing vocabulary knowledge*. Cambridge: Cambridge University Press, 1–32.

Dalton-Puffer, Christiane; Bauer-Marschallinger, Silvia; Brückl-Mackey, Katharina; Hofmann, Victoria; Hopf, Jennifer; Kröss, Lisa & Lechner, Lisa. 2018. Cognitive discourse functions in Austrian CLIL lessons: Towards an empirical validation of the CDF Construct. *European Journal of Applied Linguistics* 6(1), 5–29.

Dalton-Puffer, Christiane & Smit, Ute. 2016. Content and language integrated learning and ELF. In Marie-Luise Pitzl & Ruth Osimk-Teasdale (eds.), *English as a lingua franca*: *Perspectives and prospects. Contributions in honour of Barbara Seidlhofer*. Berlin: De Gruyter Mouton, 235–244.

Danan, Martine. 2004. Captioning and subtitling: Undervalued language learning strategies. *Meta* 49(1), 67–77.

Dang, Thi N. Y. & Webb, Stuart. 2014. The lexical profile of academic spoken English. *English for Specific Purposes* 33, 66–76.

De Cillia, Rudolf. 1997. "I glaub, daß es schon richtig ist, daß der österreichische Dialekt do muaß i sogn, holt bleibt" - Einstellungen der ÖsterreicherInnen zu ihrem Deutsch ["I think that it's right that the Austrian dialect, I have to say, just stays" - Attitudes of Austrians towards their German]. In Rudolf Muhr & Richard Schrodt (eds.), *Österreichisches Deutsch und andere nationale Varietäten plurizentrischer Sprachen in Europa.* [Austrian German and other national varieties of pluricentric languages in Europe]. Wien: Hölder-Pichler-Tempsky, 116–127.

De Cillia, Rudolf. 2003. Braucht Österreich eine Sprachenpolitik? In Brigitta Busch & Rudolf De Cillia (eds.), *Sprachenpolitik in Österreich*: *Eine Bestandsaufnahme.* [Language policy in Austria: An inventory]. Frankfurt am Main: Peter Lang, 9–42.

De Cillia, Rudolf & Haller, Michaela. 2013. Englisch und...? Vorschulisches und schulisches Sprachenlernen in Österreich [English and...? Foreign language learning at and before school in Austria]. In Rudolf De Cillia & Eva Vetter (eds.), *Sprachenpolitik in Österreich*:

Bestandsaufnahme 2011. [Language policy in Austria: Review 2011]. Frankfurt am Main: Peter Lang, 142–174.

De Cillia, Rudolf & Krumm, Hans-Jürgen. 2010. Fremdsprachenunterricht in Österreich [Foreign language teaching in Austria]. *Sociolinguistica* 24, 153–169.

De Cillia, Rudolf & Vetter, Eva (eds.). 2013. *Sprachpolitik in Österreich*: *Bestandsaufnahme 2011* [Language policy in Austria: Review 2011]. Frankfurt am Main: Peter Lang.

De Groot, Annette M. B. 2006. Effects of stimulus characteristics and background music on foreign language vocabulary learning and forgetting. *Language Learning* 56(3), 463–506.

De Wilde, Vanessa; Brysbaert, Marc & Eyckmans, June. 2019. Learning English through out-of-school exposure: Which levels of language proficiency are attained and which types of input are important? *Bilingualism: Language and Cognition* 22, 1–15.

De Wilde, Vanessa & Eyckmans, June. 2017. Game on! Young learners' incidental language learning of English prior to instruction. *Studies in Second Language Learning and Teaching* 7(4), 673–694.

DeHaan, Jonathan; Reed, W. M. & Kuwada, Katsuko. 2010. The effect of interactivity with a music video game on second language vocabulary recall. *Language Learning & Technology* 14(2), 74–94.

DeKeyser, Robert. 2003. Implicit and explicit learning. In Catherine J. Doughty & Michael H. Long (eds.), *The handbook of second language acquisition*. Malden: Blackwell, 313–348.

Dewaele, Jean-Marc. 2018. Why the dichotomy 'L1 Versus LX User' is better than 'native versus non-native speaker. *Applied Linguistics* 33(39), 236-240.

Doaee, Maryam T.; Sarkeshikian, Seyyed A. H. & Tabatabaee, Seyyed A.-M. 2017. Investigating the reliability and factor structure of the self-regulating capacity in vocabulary learning (SRCvoc) in Iranian EFL context. *The Journal of English Language Pedagogy and Practice* 10(20), 169–186.

Dóczi, Brigitta & Kormos, Judit. 2016. *Longitudinal developments in vocabulary knowledge and lexical organization*. Oxford: Oxford University Press.

Dohmen, Günther. 2001. *Das informelle Lernen*: *Die internationale Erschließung einer bisher vernachlässigten Grundform menschlichen Lernens für das lebenslange Lernen aller* [Informal learning: the international exploration of a previously neglected basic form of human learning]. Bonn: BMBF. http://www.werkstatt-frankfurt.de/fileadmin/Frankfurter_Weg/Fachtagung/BMBF_Das_informelle_Lernen.pdf (8 September, 2017).

Donzelli, Giovanna. 2007. Foreign language learners: Words they hear and words they learn. A case study. *Estudios de lingüística inglesa aplicada* 7, 103–125.

Döring, Nicola & Bortz, Jürgen. 2016. *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* [Research methods and evaluation for social scientists], 5th edn. Berlin: Springer.

Dörnyei, Zoltán. 2005. *The psychology of the language learner*: *Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Dörnyei, Zoltán. 2007. *Research methods in applied linguistics*: *Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.

Dörnyei, Zoltán. 2010. *Questionnaires in second language research*: *Construction, administration, and processing*, 2nd edn. New York: Routledge.

Dörnyei, Zoltán & Ryan, Stephen. 2015. *The psychology of the language learner revisited*. Ney York: Routledge.

Dörnyei, Zoltán & Skehan, Peter. 2003. Indiviudal differences in second language learning. In Catherine J. Doughty & Michael H. Long (eds.), *The handbook of second language acquisition*. Malden: Blackwell, 589–630.

References

Dörnyei, Zoltán & Ushioda, Ema (eds.). 2009. *Motivation, language identity and the L2 self.* Bristol: Multilingual Matters.

Dorostkar, Niku. 2014. *(Mehr-)Sprachigkeit und Lingualismus*: *Die diskursive Konstruktion von Sprache im Kontext nationaler und supranationaler Sprachenpolitik am Beispiel Österreichs* [(Multi-)Lingualism: The discursive construction of language in the context of national and supranational language policy exemplified by Austria]. Göttingen: V&R unipress.

Doughty, Catherine J. & Long, Michael H. 2003. The scope of inquiry and goals of SLA. In Catherine J. Doughty & Michael H. Long (eds.), *The handbook of second language acquisition*. Malden: Blackwell, 3–16.

Dressman, Mark; Lee, Ju S. & Sabaoui, Mohamed A. 2016. Path to English in Korea: Policies, practices, and outcomes. *English Language Teaching* 28(1), 67–78.

Drotner, Kirsten. 2008. Informal learning and digital media: Perceptions, practices and perspectives. In Kirsten Drotner, Hans S. Jensen & Kim C. Schrøder (eds.), *Informal learning and digital media*. Newcastle: Cambridge Scholars Publishing, 10–28.

Du Bois-Reymond, Manuela. 2010. Kindheit und Jugend in Europa [Childhood and youth in Europe]. In Heinz-Hermann Krüger & Cathleen Grunert (eds.), *Handbuch Kindheits- und Jugendforschung.* [Handbook childhood and youth research], 2nd edn. Wiesbaden: VS Verlag für Sozialwissenschaften, 399–418.

Du Bois-Reymond, Manuela & Chisholm, Lynne. 2006. Young Europeans in a changing world. *New directions for youth development* 113, 1–9.

Dunn, Lloyd M. & Dunn, Douglas M. 2007. *Peabody picture vocabulary test (PPVT-4)*, 4th edn. Minneapolis, MN: Pearson Assessments.

Düx, Wiebken & Sass, Erich. 2005. Lernen in informellen Kontexten: Lernpotenziale in Settings des freiwilligen Engagements [Learning in informal contexts: Learning potentials in settings of voluntary engagement]. *Zeitschrift für Erziehungswissenschaft* 8(3), 394–411.

d'Ydewalle, Géry & Pavakanun, Ubolwanna. 1995. Acquisition of a second/foreign language by viewing a television program. In Peter Winterhoff-Spurk (ed.), *Psychology of media in Europe*: *The state of the art - perspectives for the future*. Opladen: Westdeutscher Verlag, 51–64.

d'Ydewalle, Géry & Van de Poel, Marijke. 1999. Incidental foreign language acquisition by children watching subtitled television programs. *Journal of Psycholinguistic Research* 28(3), 227–244.

Eckerth, Johannes & Tavakoli, Parveneh. 2012. The effects of word exposure frequency and elaboration of word processing on incidental L2 vocabulary acquisition through reading. *Language Teaching Research* 16(2), 227–252.

Education Group GmbH. 2017a. *Oö. Jugend-Medien-Studie 2017*: *Das Medienverhalten der 11- bis 18-Jährigen* [Upper Austrian youth media study 2017: Media behaviour among 11- to 18-year-olds]. Linz. https://www.edugroup.at/detail/5-ooe-jugend-medien-studie-2017.html (12 July, 2017).

Education Group GmbH. 2017b. *Oö. Jugend-Medien-Studie 2017*: *Medienverhalten der Jugendlichen aus dem Blickwinkel der Eltern* [Upper Austrian youth media study 2017: Adolescents' media use from parents' point of view]. Linz. https://www.edugroup.at/detail/5-ooe-jugend-medien-studie-2017.html (12 July, 2017).

Education Group GmbH. 2017c. *Oö. Jugend-Medien-Studie 2017*: *Medienverhalten der Jugendlichen aus dem Blickwinkel der Jugendlichen* [Upper Austrian youth media study 2017: Adolescents' media use from their own point of view]. Linz. https://www.edugroup.at/detail/5-ooe-jugend-medien-studie-2017.html (12 July, 2017).

Edwards, Alison. 2016. *English in the Netherlands*: *Functions, forms and attitudes*. Amsterdam: John Benjamins.

Ekşi, Gül & Aydin, Hasan. 2013. What are the students doing 'out' there? An investigation of out-of-class language learning activities. *AİBÜ Sosyal Bilimler Enstitüsü Dergisi* 13(2), 191–210.

Elgort, Irina. 2011. Deliberate learning and vocabulary acquisition in a second language. *Language Learning* 61(2), 367–413.

Elgort, Irina. 2013. Effects of L1 definitions and cognate status of test items on the Vocabulary Size Test. *Language Testing* 30(2), 253–272.

Elgort, Irina. 2017. Incorrect inferences and contextual word learning in English as a second language. *Journal of the European Second Language Association* 1(1), 1–11.

Elgort, Irina. 2018. Technology-mediated second language vocabulary development: A review of trends in research methodology. *CALICO Journal* 35(1), 1–29.

Elgort, Irina; Brysbaert, Marc; Stevens, Michaël & Van Assche, Eva. 2018. Contextual word learning during reading in a second language: An eye-movement study. *Studies in Second Language Acquisition* 40(2), 341–366.

Elgort, Irina & Nation, I.S.P. 2010. Vocabulary learning in a second language: Familiar answers to new questions. In Paul Seedhouse, Steve Walsh & Chris Jenks (eds.), *Conceptualising 'learning' in applied linguistics*. Houndsmill, Basingstoke: Palgrave Macmillan, 89–104.

Elgort, Irina & Warren, Paul. 2014. L2 vocabulary learning from reading: Explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning* 64(2), 365–414.

Elley, Warwick B. 1989. Vocabulary acquisition from listening to stories. *Reading Research Quarterly* 24(2), 174–187.

Ellis, Nick C. 1994a. Consciousness in second language learning: Psychological perspectives on the role of conscious processes in vocabulary acquisition. *AILA Review* 11, 37–56.

Ellis, Nick C. 1994b. Introduction: Implicit and explicit learning - an overview. In Nick C. Ellis (ed.), *Implicit and explicit learning of languages*. London: Academic Press, 1–31.

Ellis, Nick C. 2002. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24, 143–188.

Ellis, Nick C. & Beaton, Alan. 1993. Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning* 43(4), 559–617.

Ellis, Rod. 1994. *The study of second language acquisition*. Oxford: Oxford University Press.

Ellis, Rod. 1995. Modified oral input and the acquisition of word meanings. *Applied Linguistics* 16(4), 409–441.

Ellis, Rod. 1999. Factors in the incidental acquisition of second language vocabulary from oral input. In Rod Ellis (ed.), *Learning a second language through interaction*. Amsterdam: John Benjamins, 35–61.

Ellis, Rod. 2015. *Understanding second language acquisition*, 2nd edn. Oxford: Oxford University Press.

Ellis, Rod & He, Xien. 1999. The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition* 21, 285–301.

Ellis, Rod & Heimbach, Rick. 1997. Bugs and birds: Children's acquisition of second language vocabulary through interaction. *System* 25(2), 247–259.

Ellis, Rod; Tanaka, Yoshihiro & Yamazaki, Asako. 1994. Classroom interaction, comprehension and the acquisition of L2 word meanings. *Language Learning* 44(3), 449–491.

References

Ender, Andrea. 2016. Implicit and explicit cognitive processes in incidental vocabulary acquisition. *Applied Linguistics* 37(4), 536–560.

Eskildsen, Søren W. 2018. 'We're learning a lot of new words': Encountering new L2 vocabulary outside of class. *The Modern Language Journal* 102(Supplement 2018), 46–63.

Eskildsen, Søren W. & Cadierno, Teresa. 2015. Advancing usage-based approaches to L2 studies. In Teresa Cadierno & Søren W. Eskildsen (eds.), *Usage-based perspectives on second language learning*. Berlin: De Gruyter Mouton, 1–15.

Eskildsen, Søren W. & Theodórsdóttir, GuÐrún. 2017. Constructing L2 learning spaces: Ways to achieve learning inside and outside the classroom. *Applied Linguistics* 38(2), 143–164.

European Commission. 1995. *White paper on education and training*: *Teaching and learning. Towards the learning society.* Brussels. https://publications.europa.eu/en/publication-detail/-/publication/d0a8aa7a-5311-4eee-904c-98fa541108d8/language-en (15 May, 2018).

European Commission. 2000. *A memorandum on lifelong learning.* Commission staff working paper. Brussels. https://uil.unesco.org/document/european-communities-memorandum-lifelong-learning-issued-2000 (16 July, 2018).

European Commission. 2012a. *Europeans and their languages*: *Factsheet Austria.* Special Eurobarometer 386. Brussels. http://ec.europa.eu/public_opinion/archives/eb_special_399_380_en.htm (15 October, 2014).

European Commission. 2012b. *Europeans and their languages*: *Special Eurobarometer 386 (Report).* Brussels. http://ec.europa.eu/public_opinion/index_en.htm (12 July, 2017).

European Commission/EACEA/Eurydice. 2017. *Key data on teaching languages at school in Europe – 2017 edition*: *Eurydice report.* Brussels. https://eacea.ec.europa.eu/national-policies/eurydice/content/key-data-teaching-languages-school-europe-%E2%80%93-2017-edition_en (8 May, 2018).

European Commission/EACEA/Eurydice. 2018. *Austria: Overview.* Brussels. https://eacea.ec.europa.eu/national-policies/eurydice/content/austria_en (8 May, 2018).

Evans, Sarah; Davis, Katie; Evans, Abigail; Campbell, Julie A.; Randall, David P.; Yin, Kodlee & Aragon, Cecilia. 2017. More than peer production: Fanfiction communities as sites of distributed mentoring. In Charlotte P. Lee, Steve Poltrock, Louise Barkhuus, Marcos Borges & Wendy Kellogg (eds.), *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing (CSCW '17)*. New York, NY: ACM Press, 259–272.

Eyckmans, June. 2004. *Measuring receptive vocabulary size*: *Reliability and validity of the Yes/No vocabulary test for French-speaking learners of Dutch.* Utrecht: LOT.

Eyckmans, June; Van de Velde, Hans; Van Hout, Roeland & Boers, Frank. 2007. Learners' response behaviour in yes/no vocabulary tests. In Helmut Daller, James Milton & Jeanine Treffers-Daller (eds.), *Modelling and assessing vocabulary knowledge*. Cambridge: Cambridge University Press, 59–76.

Fairclough, Marta. 2011. Testing the lexical recognition task with Spanish/English bilinguals in the United States. *Language Testing* 28(2), 273–297.

Feierabend, Sabine; Plankenhorn, Theresa & Rathgeb, Thomas. 2017. *JIM 2017 Jugend, Information, (Multi-)Media*: *Basisstudie zum Medienumgang 12- bis 19-Jähriger in Deutschland* [JIM 2017 youth, information, (multi)media: Basic study of media use among 12- to 19-year-olds]. Stuttgart: Medienpädagogischer Forschungsverbund Südwest (mpfs). https://www.mpfs.de/studien/jim-studie/2017/ (6 June, 2018).

Feitelson, Dina; Goldstein, Zahava; Iraqi, Jihad & Share, David L. 1993. Effects of listening to story reading on aspects of literacy acquisition in a diglossic situation. *Reading Research Quarterly* 28(1), 70–79.

Field, Andy; Miles, Jeremy & Field, Zoë. 2012. *Discovering statistics using R*. Thousand Oaks, CA: SAGE.

References

Fischer, Gero & Doleschal, Ursula. 2013. Von Minderheitensprachen zu Nachbarsprachen - Die Rolle der Minderheitensprachen in Österreichs Bildungswesen 2011 [From minority languages to neighbouring languages - The role of minority languages in the Austrian education system 2011]. In Rudolf De Cillia & Eva Vetter (eds.), *Sprachenpolitik in Österreich*: *Bestandsaufnahme 2011*. [Language policy in Austria: Review 2011]. Frankfurt am Main: Peter Lang, 68–93.

Fitzpatrick, Tess. 2007. Productive vocabulary tests and the search for concurrent validity. In Helmut Daller, James Milton & Jeanine Treffers-Daller (eds.), *Modelling and assessing vocabulary knowledge*. Cambridge: Cambridge University Press, 116–132.

Fitzpatrick, Tess. 2012. Tracking the changes: Vocabulary acquisition in the study abroad context. *The Language Learning Journal* 40(1), 81–98.

Fitzpatrick, Tess & Clenton, John. 2010. The challenge of validation: Assessing the performance of a test of productive vocabulary. *Language Testing* 27(4), 537–554.

Fitzpatrick, Tess & Clenton, Jon. 2017. Making sense of learner performance on tests of productive vocabulary knowledge. *TESOL Quarterly* 51(4), 844–867.

Fitzpatrick, Tess & Meara, Paul. 2004. Exploring the validity of a test of productive vocabulary. *VIAL (Vigo International Journal of Applied Linguistics)* 1, 55–73.

Flammer, August & Schaffner, Brigitta. 2003. Adolescent leisure across European nations. *New directions for child and adolescent development* 99, 65–77.

Fleck, Elfie. 2013. Zur Situation von lebensweltlich mehrsprachigen SchülerInnen: Aktuelle Lage und neuere Entwicklungen in der Bildungspolitik [On the situation of multilingual students: Current situation and recent developments in educational policy]. In Rudolf De Cillia & Eva Vetter (eds.), *Sprachenpolitik in Österreich*: *Bestandsaufnahme 2011*. [Language policy in Austria: Review 2011]. Frankfurt am Main: Peter Lang, 9–28.

Francis, Winthrop N. & Kučera, Henry. 1982. *Frequency analysis of English usage*: *Lexicon and grammar*. Boston: Houghton Mifflin.

Freed, Barbara F.; Segalowitz, Norman & Dewey, Dan P. 2004. Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition* 26(2), 275–301.

Freeman, Mike. 1999. The language learning activities of students of EFL and French at two universities. *The Language Learning Journal* 19(1), 80–88.

Furlong, Andy (ed.). 2017. *Routledge handbook of youth and young aduulthood*, 2nd edn. London: Routledge.

Gadermann, Anne M.; Guhn, Martin & Zumbo, Bruno D. 2012. Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide - practical assessment, research & evaluation. *Practical Assessment, Research & Evaluation* 17(3), 1–12.

Gałecki, Andrzej & Burzykowski, Tomasz. 2013. *Linear mixed-effects models using R*: *A step-by-step approach*. New York: Springer.

Ganzeboom, Harry B.G. 2010. *Harry Ganzeboom's tools for deriving occupational status measures from ISCO-08 with interpretative notes to ISCO-08*. http://www.harryganzeboom.nl/isco08/index.htm (17 May, 2016).

Ganzeboom, Harry B.G. & Treiman, Donald J. 2010. *International stratification and mobility file*: *Conversion tools* . Amsterdam: Department of Social Research Methodology. http://www.harryganzeboom.nl/isco08/index.htm.

Gardner, Dee. 2007. Validating the construct of word in applied corpus-based vocabulary research: A critical survey. *Applied Linguistics* 28(2), 241–265.

Garnier, Mélodie & Schmitt, Norbert. 2015. The PHaVE List: A pedagogical list of phrasal verbs and their most frequent meaning senses. *Language Teaching Research* 19(6), 645–666.

Gass, Susan. 1999. Discussion: Incidental vocabulary learning. *Studies in Second Language Acquisition* 21, 319–333.

Gerngross, Günter; Puchta, Herbert; Holzmann, Christian; Stranks, Jeff & Lewis-Jones, Peter. 2008a. *More! 2*: *Student's book/Workbook*. Innsbruck: Helbling Languages.

Gerngross, Günter; Puchta, Herbert; Holzmann, Christian; Stranks, Jeff & Lewis-Jones, Peter. 2008b. *More! 3*: *Student's book/Workbook*. Innsbruck: Helbling Languages.

Gerngross, Günter; Puchta, Herbert; Holzmann, Christian; Stranks, Jeff & Lewis-Jones, Peter. 2009a. *More! 1*: *Student's book/Workbook*, 2nd edn. Innsbruck: Helbling Languages.

Gerngross, Günter; Puchta, Herbert; Holzmann, Christian; Stranks, Jeff & Lewis-Jones, Peter. 2009b. *More! 4*: *Student's book/Workbook*. Innsbruck: Helbling Languages.

Gesetz zum Schutz der Jugend: (Wiener Jugendschutzgesetz 2002) [Law for the protection of youth (Viennese youth protection law 2002)]. In *LGBl. Nr. 08/2007 as amended by LGBl. Nr. 10/2013*.

Gleason, Benjamin. 2015. New literacies practices of teenage Twitter users. *Learning, Media and Technology* 41(1), 31–54.

Gnutzmann, Claus & Intemann, Frauke. 2005. Introduction: The globalisation of English. Language, politics and the English language classroom. In Claus Gnutzmann & Frauke Intemann (eds.), *The globalisation of English and the English language classroom.* Tübingen: Gunter Narr Verlag, 9–24.

Godfroid, Aline. 2019. Sensitive measures of vocabulary knowledge and processing: Expanding Nation's framework. In Stuart A. Webb (ed.), *The Routledge handbook of vocabulary studies*. Milton Park, Abingdon: Routledge, 433–453.

Godwin-Jones, Robert. 2018. Contextualized vocabulary learning. *Language Learning & Technology* 22(3), 1–19.

Gohar, Manoochehr J.; Rahmanian, Mahboubeh & Soleimani, Hassan. 2018. Technique Feature Analysis or Involvement Load Hypothesis: Estimating their predictive power in vocabulary learning. *Journal of Psycholinguistic Research* 47(4), 859–869.

González-Fernández, Beatriz. 2019. Conceptualising vocabulary knowledge: An empirical examination across L2 learner populations and proficiency levels (Paper presented at Vocab@Leuven 2019. Leuven, 3 July 2019).

González-Fernández, Beatriz & Schmitt, Norbert. 2015. How much collocation knowledge do L2 learners have? The effects of frequency and amount of exposure. *ITL - International Journal of Applied Linguistics* 166(1), 94–126.

González-Fernández, Beatriz & Schmitt, Norbert. 2019. Word knowledge: Exploring the relationships and order of acquisition of vocabulary knowledge components. *Applied Linguistics* 52, 1–26.

Goulden, Robin; Nation, I.S.P. & Read, John. 1990. How large can a receptive vocabulary be? *Applied Linguistics* 11(4), 341–363.

Graddol, David. 2006. *English next*: *Why global English may mean the end of 'English as a foreign language':* British Council. https://englishagenda.britishcouncil.org/continuing-professional-development/cpd-researchers/english-next (14 August, 2018).

Granena, Gisela. 2013. Cognitive aptitudes for second language learning and the LLAMA language aptitude test. In Gisela Granena & Michael H. Long (eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment*. Amsterdam, Philadelphia: John Benjamins Publishing Company, 105–129.

Grau, Maike. 2009. Worlds apart? English in German youth cultures and in educational settings. *World Englishes* 28(2), 160–174.

Greene, Jennifer C.; Caracelli, Valerie J. & Graham, Wendy F. 1989. Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis* 11(3), 255–274.

Großegger, Beate. 2014. *Next Generation*: *Jugend zwischen Problemfall und Hoffnungsträger?* [Next Generation: Youth between problem cases and new hopes?]. Wien: Institut für Jugendkulturforschung. https://jugendkultur.at/publikationen/online/.

Großegger, Beate. 2017. Zwischen Freakout und Normcore: Jugend und Jugendkulturen in den späten 2010er Jahren. *ÖRF* 25(1), 7–16.

Gu, Yongqi. 2013. Vocabulary learning strategies. In Carol A. Chapelle (ed.), *The encyclopedia of applied linguistics*. Oxford, UK: Blackwell, n.d.

Gu, Yongqi & Johnson, Robert K. 1996. Vocabulary learning strategies and language learning outcomes. *Language Learning* 46(4), 643–679.

Gyllstad, Henrik; Vilkaitė, Laura & Schmitt, Norbert. 2015. Assessing vocabulary size through multiple-choice formats: Issues with guessing and sampling rates. *ITL - International Journal of Applied Linguistics* 166(2), 278–306.

Haastrup, Kirsten. 1991. *Lexical inferencing procedures or talking about words*: *Receptive procedures in foreign language learning with special reference to English*. Tübingen: Gunter Narr Verlag.

Haastrup, Kirsten. 2008. Lexical inferencing procedures in two languages. In Dorte Albrechtsen, Kirsten Haastrup & Birgit Henriksen (eds.), *Vocabulary and writing in a first and second language*. London: Palgrave Macmillan, 67–111.

Hahn, Magdalena. 2017. *Extramural vocabulary acquisition*: *A survey of students in Viennese vocational business middle schools (HAS)*. Wien: University of Vienna, MA thesis.

Hahn, Magdalena. 2018. Extramural vocabulary acquisition: A survey of students in Viennese vocational business middle schools (HAS). *Vienna English Working Papers (VIEWS)* 27, 1–21.

Hannibal Jensen, Signe. 2017. Gaming as an English language learning resource among young children in Denmark. *CALICO Journal* 34(1), 1–19.

Hannibal Jensen, Signe. 2019. Language learning in the wild: A young user perspective. *Language Learning & Technology* 23(1), 72–86.

Hansson, Åse & Gustafsson, Jan-Eric. 2013. Measurement invariance of socioeconomic status across migrational background. *Scandinavian Journal of Educational Research* 57(2), 148–166.

Harrell, Joanne S.; Bradley, Chyrise; Dennis, Jennifer; Frauman, Annette C. & Criswell, Elaine S. 2000. School-based research: Problems of access and consent. *Journal of Pediatric Nursing* 15(1), 14–21.

Harring, Marius; Witte, Matthias D. & Burger, Timo (eds.). 2016. *Handbuch informelles Lernen*: *Interdisziplinäre und internationale Perspektiven* [Handbook informal learning: Interdisciplinary and international perspectives]. Weinheim: Beltz Juventa.

Harrington, Michael. 2018. *Lexical facility*: *Size, recognition speed and consistency as dimensions of second language vocabulary knowledge*. London: Palgrave Macmillan.

Harrington, Michael & Carey, Michael. 2009. The on-line Yes/No test as a placement tool. *System* 37(4), 614–626.

Harsch, Claudia & Hartig, Johannes. 2016. Comparing C-tests and Yes/No vocabulary size tests as predictors of receptive language skills. *Language Testing* 33(4), 555–575.

Hasebrink, Uwe. 2007. English, youth and media environments. In Margie Berns, Kees De Bot & Uwe Hasebrink (eds.), *In the presence of English*: *Media and European youth*. New York: Springer, 89–110.

Hashemi, Mohammad R. & Babaii, Esmat. 2013. Mixed methods research: Toward new research designs in applied linguistics. *The Modern Language Journal* 97(4), 828–852.

Haß, Frank (ed.). 2007. *Red Line 2*: *Coursebook/Workbook*. Wien: Öbv.

Haß, Frank (ed.). 2008a. *Red Line 1*: *Coursebook/Workbook*. Wien: Öbv.

Haß, Frank (ed.). 2008b. *Red Line 3*: *Coursebook/Workbook*. Wien: Öbv.

Haß, Frank (ed.). 2009. *Red Line 4*: *Coursebook/Workbook*. Wien: Öbv.

Hellmayr, Georg; Waba, Stephan & Mlakar, Heike. 2010a. *Prime Time 5*: *Coursebook*, 2nd edn. Wien: Öbv.

Hellmayr, Georg; Waba, Stephan & Mlakar, Heike. 2010b. *Prime Time 6*: *Coursebook*. Wien: Öbv.

Henriksen, Birgit. 1999. Three dimensions of vocabulary development. *Studies in Second Language Acquisition* 21, 303–317.

Henriksen, Birgit. 2008. Declarative lexical knowledge. In Dorte Albrechtsen, Kirsten Haastrup & Birgit Henriksen (eds.), *Vocabulary and writing in a first and second language*. London: Palgrave Macmillan, 22–66.

Henry, Alastair. 2014. Swedish students' beliefs about learning English in and outside of school. In David Lasagabaster, Aintzane Doiz & Juan M. Sierra (eds.), *Motivation and foreign language learning*: *From theory to practice*. Amsterdam: John Benjamins Pub. Co, 93–116.

Hirsh, David & Nation, I.S.P. 1992. What vocabulary is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language* 8(2), 689–696.

Hoffmann, Charlotte. 2000. The spread of English and the growth of multilingualism with English in Europe. In Jasone Cenoz & Ulrike Jessner (eds.), *English in Europe*: *The acquisition of a third language*. Clevedon: Multilingual Matters, 1–37.

Horst, Marlise. 2010. How well does teacher talk support incidental vocabulary acquisition? *Reading in a Foreign Language* 22(1), 161–180.

Horst, Marlise; Cobb, Tom & Meara, Paul. 1998. Beyond a Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language* 11(2), 207–223.

Howell, David C. 2013. *Statistical methods for psychology*, 8th edn. Belmont, CA: Wadsworth Cengage Learning.

Hu, Hsueh-chao M. & Nassaji, Hossein. 2016. Effective vocabulary learning tasks: Involvement Load Hypothesis versus Technique Feature Analysis. *System* 56, 28–39.

Hu, Hsueh-chao M. & Nation, I.S.P. 2000. Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language* 13(1), 403–430.

Huber-Bachmann, Eva. 2017. *Österreich*: *Zahlen, Daten, Fakten 16/17* [Austria: Figures, data, facts 16/17], 12th edn. Wien: Statistik Austria.

Huckin, Thomas N. & Coady, James. 1999. Incidental vocabulary acquisition in a second language: A review. *Studies in Second Language Acquisition* 21, 181–193.

Huibregtse, Ineke; Admiraal, Wilfried & Meara, Paul. 2002. Scores on a yes-no vocabulary test: Correction for guessing and response style. *Language Testing* 19(3), 227–245.

Hulstijn, Jan H. 2003. Incidental and intentional learning. In Catherine J. Doughty & Michael H. Long (eds.), *The handbook of second language acquisition*. Malden: Blackwell, 348–381.

Hulstijn, Jan H. 2007. The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal* 91(4), 663–667.

Hulstijn, Jan H. 2013. Incidental learning in second language acquisition. In Carol A. Chapelle (ed.), *The encyclopedia of applied linguistics*. Oxford, UK: Blackwell, n.d.

Hulstijn, Jan H. & Laufer, Batia. 2001. Some empirical evidence for the Involvement Load Hypothesis in vocabulary acquisition. *Language Learning* 51(3), 539–558.

Hyland, Fiona. 2004. Learning autonomously: Contextualising out-of-class English language learning. *Language Awareness* 13(3), 180–202.

IBM Corp. 2016. *IBM SPSS Statistics for Windows* (22.0). Armonk, NY: IBM Corp.

Inaba, Miho. 2019. *Second language literacy practices and language learning outside the classroom*. Bristol: Multilingual Matters.

Ingvarsdóttir, Hafdís & Jóhannsdóttir, Ásrún. 2017. Learning and using English: The views of learners at the end of compulsory education. In Birna Arnbjörnsdóttir & Hafdís Ingvarsdóttir (eds.), *Language development across the life span*: *The impact of English on education and work in Iceland*. Cham: Springer, 79–94.

Inozu, Julide; Sahinkarakas, Sehnaz & Yumru, Hulya. 2010. The nature of language learning experiences beyond the classroom and its learning outcomes. *US-China Foreign Language* 8(1), 14–21.

Institut für Jugendkulturforschung [Institute for Research on Youth Culture]. 2018. *Studie: Jugend und digitale Medien*: *Neue Studie zu digitalen Lebenswelten junger ÖsterreicherInnen* [Study: Youth and digital media: New study on digital lifeworlds of young Austrians]: Wien. https://jugendkultur.at/studie-jugend-digitale-medien/#more-11523 (6 June, 2018).

Integral Markt- und Meinungsforschung Integral market and opinion research. 2017a. *Austrian Internet Monitor*: *Kommunikation und IT in Österreich, 1. Quartal 2017* [Austrian Internet Monitor: Communication and IT in Austria, 1st quarter 2017]. Vienna. http://www.integral.co.at/de/downloads/ (17 May, 2018).

Integral Markt- und Meinungsforschung Integral market and opinion research. 2017b. *Austrian Internet Monitor*: *Kommunikation und IT in Österreich, 3. Quartal 2017* [Austrian Internet Monitor: Communication and IT in Austria, 3rd quarter 2017]. Vienna. http://www.integral.co.at/de/downloads/ (17 May, 2018).

Interactive Advertising Bureau (IAB) Austria. 2017. *IAB Trendmonitor*: *Zahlungsbereitschaft für Content* [Willingness to pay for content]. Vienna. https://www.iab-austria.at/iab-trendmonitor-zahlungsbereitschaft-fuer-content/ (12 July, 2017).

Interactive Advertising Bureau (IAB) Austria; Bundesverband Digitale Wirtschaft (BVDW) & Interactive Advertising Bureau (IAB) Switzerland. 2016. *DACH Mediennutzung '16*: *Gemeinschaftsstudie von IAB Austria, BVDW, IAB Switzerland* [DACH Media use '16: Collaborative study by IAB Austria, BVDW, IAB Switzerland]. https://www.iab-austria.at/wp-content/uploads/2017/02/DACH_Mediennutzung_Launch_2016.pdf (12 July, 2017).

International Labour Office. 2012. *International standard classification of occupations (ISCO-08)*: *Structure, group definitions and correspondence tables.* Geneva. https://www.ilo.org/public/english/bureau/stat/isco/isco08/ (17 May, 2016).

Isbell, Daniel R. 2018. Online informal language learning: Insights from a Korean learning community. *Language Learning & Technology* 22(3), 82–102.

Ishikawa, Shin'ichiro; Uemura, Toshihiko; Kaneda, M.; Shmizu S.; Sugimori, N.; Tono, Yukio; Mochizuki, Masamichi & Murata, M. 2003. *JACET 8000*: *JACET list of 8000 basic words*. Tokyo: Japan Association of College English Teachers (JACET).

Ito, Mizuko; Baumer, Sonja; Bittanti, Matteo; Boyd, Danah; Cody, Rachel; Herr-Stephenson, Becky; Horst, Heather A.; Lange, Patricia G.; Mahendran, Dilan; Martínez, Katynka Z.; Pascoe, C. J.; Perkel, Dan; Robinson, Laura; Sims, Christo & Tripp, Lisa. 2010. *Hanging out, messing around, and geeking out*: *Kids living and learning with new media*. Cambridge, MA: MIT Press.

Ivankova, Nataliya V. 2013. Implementing quality criteria in designing and conducting a sequential QUAN -> QUAL mixed methods study of student engagement with learning applied research methods online. *Journal of Mixed Methods Research* 8(1), 25–51.

References

Ivankova, Nataliya V.; Creswell, John W. & Stick, Sheldon L. 2006. Using mixed-methods sequential explanatory design: From theory to practice. *Field Methods* 18(1), 3–20.

Jabbari, Nasser & Eslami, Zohreh R. 2019. Second language learning in the context of massively multiplayer online games: A scoping review. *ReCALL* 31(1), 92–113.

Jakonen, Teppo. 2014. Building bridges - How secondary school pupils bring their informal learning experiences into a content and language integrated (CLIL) classroom. *APPLES – Journal of Applied Language Studies* 8(1), 7–28.

Jarvis, Huw & Achilleos, Marianna. 2013. From computer assisted language learning (CALL) to mobile assisted language use (MALU). *The Electronic Journal for English as a Second Language (TESTL-EJ)* 16(4), 1–18.

Jarvis, Scott. 2013. Defining and measuring lexical diversity. In Scott Jarvis & Helmut Daller (eds.), *Vocabulary knowledge*: *Human ratings and automated measures*. Amsterdam: John Benjamins, 13–43.

Jeeves, Anna. 2017. Perceptions of relevance of English education at secondary school. In Birna Arnbjörnsdóttir & Hafdís Ingvarsdóttir (eds.), *Language development across the life span*: *The impact of English on education and work in Iceland*. Cham: Springer, 113–141.

Jelani, Nurul A. M. & Boers, Frank. 2018. Examining incidental vocabulary acquisition from captioned video: Does test modality matter? *ITL - International Journal of Applied Linguistics* 169(1), 169–190.

Jiménez-Catalán, Rosa M. (ed.). 2010. *Gender perspectives on vocabulary in foreign and second languages*. Houndsmill, Basingstoke: Palgrave Macmillan.

Jiménez-Catalán, Rosa M. & Moreno Espinosa, Soraya. 2005. Using Lex30 to measure the L2 productive vocabulary of Spanish primary learners of EFL. *VIAL (Vigo International Journal of Applied Linguistics)* 2, 27–44.

Jóhannsdóttir, Ásrún. 2017. English exposure and vocabulary proficiency at the onset of English instruction. In Birna Arnbjörnsdóttir & Hafdís Ingvarsdóttir (eds.), *Language development across the life span*: *The impact of English on education and work in Iceland*. Cham: Springer, 57–78.

Johnson, R. B. & Onwuegbuzie, Anthony J. 2004. Mixed methods research: A research paradigm whose time has come. *Educational Researcher* 33(7), 14–26.

Johnsson-Smaragdi, Ulla. 2009. Vergleichende Jugendmedienforschung: Probleme und Perspektiven [Comparative youth media research: Problems and perspectives]. In Angela Schorr (ed.), *Jugendmedienforschung*: *Forschungsprogramme, Synopse, Perspektiven.* [Youth media research: Research programs, synopsis, perspectives]. Wiesbaden: VS Verlag für Sozialwissenschaften, 165–200.

Jurkovič, Violeta. 2019. Online informal learning of English through smartphones in Slovenia. *System* 80, 27–37.

Kachru, Braj B. 1985. Standards, codification and sociolinguistic realism: The English language in the outer circle. In Randolph Quirk & Henry G. Widdowson (eds.), *English in the world: Teaching and learning the language and literatures*. Cambridge: Cambridge University Press, 11–30.

Kalaja, Paula; Alanen, Riikka; Palviainen, Åsa & Dufva, Hannele. 2011. From milk cartons to English roommates: Content and agency in L2 learning beyond the classroom. In Phil Benson & Hayo Reinders (eds.), *Beyond the language classroom*. Houndsmill, Basingstoke: Palgrave Macmillan, 47–58.

Kane, Michael T. 1992. An argument-based approach to validity. *Psychological Bulletin* 112(3), 527–535.

Kane, Michael T. 2012. Validating score interpretations and uses: Messick Lecture (Language Testing Research Colloquium, Cambridge, April 2010). *Language Testing* 29(1), 3–17.

References

Karami, H. 2012. The development and validation of a bilingual version of the Vocabulary Size Test. *RELC Journal* 43(1), 53–67.

Karpa, Dietrich; Overwien, Bernd & Plessow, Oliver (eds.). 2015. *Außerschulische Lernorte in der politischen und historischen Bildung* [Learning spaces outside school in political and historical education]. Immenhausen: Prolog-Verlag.

Kaur, Naginder. 2015. Making meaning of vocabulary learning: Seizing opportunities at opportune moments. *GEMA Online® Journal of Language Studies* 15(2), 1–16.

Ke, I-Chung & Cahyani, Hilda. 2014. Learning to become users of English as a lingua franca (ELF): How ELF online communication affects Taiwanese learners' beliefs of English. *System* 46, 28–38.

Keating, Gregory D. 2008. Task effectiveness and word learning in a second language: The Involvement Load Hypothesis on trial. *Language Teaching Research* 12(3), 365–386.

Kim, Seongho. 2015. *ppcor: Partial and semi-partial (part) correlation* (R package version 1.1).

Kim, YouJin. 2008. The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning* 58(2), 285–325.

Kirby, Kris N. & Gerlanc, Daniel. 2013. BootES: An R package for bootstrap confidence intervals on effect sizes. *Behavior research methods* 45(4), 905–927.

Kirchhoff, Sabine; Kuhnt, Sonja; Lipp, Peter & Schlawin, Siegfried. 2001. *Der Fragebogen*: *Datenbasis, Konstruktion und Auswertung* [The questionnaire: Data, design and analysis]. Wiesbaden: VS Verlag für Sozialwissenschaften.

Kiss, George R.; Armstrong, Christine & Milroy, Robert. 1973. *An associative thesaurus of English*. Wakefield: EP Microfilms.

Klapper, John. 2008. Vocabulary learning strategies in independent second language learning. In Stella Hurd & Tim Lewis (eds.), *Language learning strategies in independent settings*. Bristol: Multilingual Matters, 159–178.

Klein, Hans-Peter. 2015. "Wir sind in Oswiecim gewesen und haben Auschwitz gesehen – Lernort Auschwitz" [We've been to Oswiecim and saw Auschwitz – learning space Auschwitz. In Dietrich Karpa, Bernd Overwien & Oliver Plessow (eds.), *Außerschulische Lernorte in der politischen und historischen Bildung.* [Learning spaces outside school in political and historical education]. Immenhausen: Prolog-Verlag, 142–150.

Knight, Tracey. 2007. *Beyond the classroom walls*: *A study of out of class English use by adult community college ESL students.* Portland, OR: Portland State University, MA thesis.

Koivistoinen, Hilkka A. 2014. Crossing geographies of language learning – the case of 'a successful pupil'. *Classroom Discourse* 6(1), 20–32.

Koolstra, Cees M. & Beentjes, Jonannes W. J. 1999. Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational Technology Research and Development* 47(1), 51–60.

Kral, Claudia. 2012. *Marktges(ch)ehen - Sprachregime Brunnenmarkt?*: *Ein ethnographischer Zugang zur Linguistic Landscape* [Marktge(ch)ehen - Language use at Brunnenmarkt: An ethnographic approach to the linguistic landscape]. Wien: University of Vienna, MA thesis.

Kremmel, Benjamin. 2016. Word families and frequency bands in vocabulary tests: Challenging conventions. *TESOL Quarterly* 50(4), 976–987.

Kremmel, Benjamin. 2017. *Development and initial validation of a diagnostic computer-adaptive profiler of vocabulary knowledge.* Nottingham: University of Nottingham, PhD thesis.

Kremmel, Benjamin & Schmitt, Norbert. 2016. Interpreting vocabulary test scores: What do various item formats tell us about learners' ability to employ words? *Language Assessment Quarterly* 13(4), 377–392.

Krueger, Richard A. & Casey, Mary A. 2001. Designing and conducting focus group interviews. *Social Development Papers* 36, 4–23.

References

Kuckartz, Udo. 2014. *Qualitative text analysis*: *A guide to methods, practice & using software*. Los Angeles: SAGE.

Kuckartz, Udo. 2016. *Qualitative Inhaltsanalyse*: *Methoden, Praxis, Computerunterstützung* [Qualitative content analysis: Methods, practice and software support], 3rd edn. Weinheim: Beltz Juventa.

Kuppens, An H. 2010. Incidental foreign language acquisition from media exposure. *Learning, Media and Technology* 35(1), 65–85.

Kurtz, Jürgen. 2015. Dimensionen einer fremdsprachendidaktischen Theorie der Lernorte [Dimensions of a foreign language pedagogical theory of learning spaces]. In Eva Burwitz-Melzer, Frank G. Königs & Claudia Riemer (eds.), *Lernen an allen Orten? Die Rolle der Lernorte beim Lehren und Lernen.* [Learning in all places? The role of learning spaces in teaching and learning of languages]. Tübingen: Narr Verlag, 106–116.

Kusyk, Meryl. 2017. The development of complexity, accuracy and fluency in L2 written production through informal participation in online activities. *CALICO Journal* 34(1), 75–96.

Kusyk, Meryl & Sockett, Geoffrey. 2012. From informal resource usage to incidental language acquisition: Language uptake from online television viewing in English. *ASp - la revue du GERAS*(62), 45–65.

Kuure, Leena. 2011. Places for learning: Technology-mediated language learning practices beyond the classroom. In Phil Benson & Hayo Reinders (eds.), *Beyond the language classroom.* Houndsmill, Basingstoke: Palgrave Macmillan, 35–46.

Lai, Chun. 2015. Perceiving and traversing in-class and out-of-class learning: Accounts from foreign language learners in Hong Kong. *Innovation in Language Learning and Teaching* 9(3), 265–284.

Lai, Chun & Gu, Mingyue. 2011. Self-regulated out-of-class language learning with technology. *Computer Assisted Language Learning* 24(4), 317–335.

Lai, Chun; Hu, Xiao & Lyu, Boning. 2018. Understanding the nature of learners' out-of-class language learning experience with technology. *Computer Assisted Language Learning* 31(1-2), 114–143.

Lai, Chun; Yeung, Yuk & Hu, Jingjing. 2016. University student and teacher perceptions of teacher roles in promoting autonomous language learning with technology outside the classroom. *Computer Assisted Language Learning* 29(4), 703–723.

Lai, Chun & Zheng, Dongping. 2018. Self-directed use of mobile devices for language learning beyond the classroom. *ReCALL* 30(3), 299–318.

Lai, Chun; Zhu, Weimin & Gong, Gang. 2015. Understanding the quality of out-of-class English learning. *TESOL Quarterly* 49(2), 278–308.

Lam, Wan S. E. 2000. L2 literacy and the design of the self: A case study of a teenager writing on the internet. *TESOL Quarterly* 34(3), 457–482.

Lamb, Martin. 2002. Explaining successful language learning in difficult circumstances. *Prospect* 17(2), 35–52.

Lamb, Martin. 2004a. Integrative motivation in a globalizing world. *System* 32(1), 3–19.

Lamb, Martin. 2004b. 'It depends on the students themselves': Independent language learning at an Indonesian state school. *Language, Culture and Curriculum* 17(3), 229–245.

Lancaster, Nina K. 2018. Extramural exposure and language attainment: The examination of input-related variables in CLIL programmes. *Porta Linguarum* 29, 91–114.

Landry, Rodrigue & Bourhis, Richard Y. 1997. Linguistic landscape and ethnolinguistic vitality. *Journal of Language and Social Psychology* 16(1), 23–49.

Larson-Hall, Jenifer. 2016. *A guide to doing statistics in second language research using SPSS and R*, 2nd edn. New York: Routledge.

References

Larson-Hall, Jenifer & Plonsky, Luke. 2015. Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning* 65(S1), 127–159.

Laufer, Batia. 1989. What percentage of text-lexis is essential for comprehension? In Christer Laurén & Marianne Nordman (eds.), *Special language*: *From humans thinking to thinking machines*. Clevedon: Multilingual Matters, 316–323.

Laufer, Batia. 1997a. The lexical plight in second language reading: Words you don't know, words you think you know and words you can't guess. In James Coady & Thomas N. Huckin (eds.), *Second language vocabulary acquisition*: *A rationale for pedagogy*. Cambridge: Cambridge University Press, 20–34.

Laufer, Batia. 1997b. What's in a word that makes it hard or easy: Some intralexical factors that affect learning of words. In Norbert Schmitt & Michael McCarthy (eds.), *Vocabulary*: *Description, acquisition and pedagogy*. Cambridge: Cambridge University Press, 140–155.

Laufer, Batia. 1998. The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics* 12, 255–271.

Laufer, Batia. 2000. Task effect on instructed vocabulary learning: The hypothesis of 'involvement'. In AILA '99 Tokyo Organizing Committee (ed.), *Selected papers from AILA '99 Tokyo*. Tokyo: Waseda University Press, 47–62.

Laufer, Batia. 2001. Quantitative evaluation of vocabulary: How it can be done and what it is good for. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. McNamara & K. O'Loughlin (eds.), *Experimenting with uncertainty*: *Essays is honour of Alan Davies*. Cambridge: Cambridge University Press, 241–250.

Laufer, Batia. 2005. Focus on form in second language vocabulary learning. *EUROSLA Yearbook* 5, 223–250.

Laufer, Batia & Goldstein, Zahava. 2004. Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning* 54(3), 399–436.

Laufer, Batia & Hulstijn, Jan H. 2001. Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics* 22(1), 1–26.

Laufer, Batia & Levitzky-Aviad, Tami. 2018. Loanword proportion in vocabulary size tests: Does it make a difference? *ITL - International Journal of Applied Linguistics* 169(1), 95–114.

Laufer, Batia & McLean, Stuart. 2016. Loanwords and vocabulary size test scores: A case of different estimates for different L1 learners. *Language Assessment Quarterly* 13(3), 202–217.

Laufer, Batia & Nation, I.S.P. 1995. Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics* 16(3), 307–322.

Laufer, Batia & Nation, I.S.P. 1999. A vocabulary-size test of controlled productive ability. *Language Testing* 16(1), 33–51.

Laufer, Batia & Nation, I.S.P. 2012. Vocabulary. In Susan M. Gass & Alison Mackey (eds.), *The Routledge handbook of second language acquisition*. London: Routledge, 163–176.

Laufer, Batia & Paribakht, T. S. 1998. The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning* 48(3), 365–391.

Laufer, Batia & Ravenhorst-Kalovski, Geke C. 2010. Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language* 22(1), 15–20.

Lazaraton, Anne. 2000. Current trends in research methodology and statistics in applied linguistics. *TESOL Quarterly* 34(1), 175.

Le Nguyen, Thi C. & Nation, I.S.P. 2011. A bilingual Vocabulary Size Test of English for Vietnamese learners. *RELC Journal* 42(1), 86–99.

References

Lee, Ju S. 2019a. Informal digital learning of English and second language vocabulary outcomes: Can quantity conquer quality? *British Journal of Educational Technology* 50(2), 767–778.

Lee, Ju S. 2019b. Quantity and diversity of informal digital learning of English. *Language Learning & Technology* 23(1), 114–126.

Lee, Ju S. & Dressman, Mark. 2018. When IDLE hands make an English workshop: Informal digital learning of English and language proficiency. *TESOL Quarterly* 52(2), 435–445.

Lefever, Samúel. 2010. English skills of young learners in Iceland: "I started talking English when I was 4 years old. It just bang… just fall into me". *Ráðstefnurit Netlu – Menntakvika*, 1–17.

Legutke, Michael & Thiel, Wolfgang. 1983. *Airport*: *Ein Projekt für den Englischunterricht in Jahrgangsstufe 6* [Airport: A project for English lessons in year 6]. Wiesbaden: HIBS.

Lekkai, Ina. 2014. Incidental foreign-language acquisition by children watching subtitled television programs. *TOJET (The Turkish Online Journal of Educational Technology)* 13(4), 81–87.

Lemhöfer, Kristin & Broersma, Mirjam. 2012. Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior research methods* 44(2), 325–343.

Lemke, Jay; Lecusay, Robert; Cole, Michael & Michalchik, Vera. 2015. *Documenting and accessing learning in informal and media-rich environments*. Cambridge, MA: MIT Press.

Leppänen, Sirpa; Pitkänen-Huhta, Anne; Nikula, Tarja; Kytölä, Samu; Törmäkangas, Timo; Nissinen, Kari; Kääntä, Leila; Räisänen, Tiina; Laitinen, Mikko; Koskela, Heidi; Lähdesmäki, Salla & Jousmäki, Henna. 2011. *National survey on the English language in Finland*: *Uses, meanings and attitudes* (Studies in Variation, Contacts and Change in English Paper presented at). Jyväskylä: VARIENG.

Lien, Nanna; Friestad, Christine & Klepp, Knut-Inge. 2001. Adolescents' proxy reports of parents'socioeconomic status: How valid are they? *Journal of Epidemiology & Community Health* 55, 731–737.

Lin, Lu-fang. 2010. English learners' incidental vocabulary acquisition in the video-based CALL program. *Asian EFL Journal* 12(4), n.d.

Lin, Phoebe M.S. 2014. Investigating the validity of internet television as a resource for acquiring L2 formulaic sequences. *System* 42, 164–176.

Lindgren, Eva & Muñoz, Carmen. 2013. The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism* 10(1), 105–129.

Littlewood, William & Liu, Ngan F. 1996. *Hong Kong students and their English*. Hong Kong: Hong Kong University.

Liu, Xianghu. 2014. Students' perceptions of autonomous out-of-class learning through the use of computers. *English Language Teaching* 7(4), 74–82.

Livingstone, Sonia. 2014. Developing social media literacy: How children learn to interpret risky opportunities on social network sites. *Communications: The European Journal of Communication Research* 39(3), 283–303.

Livingstone, Sonia; d'Haenens, Leen & Hasebrink, Uwe. 2001. Childhood in Europe: Contexts for comparison. In Sonia Livingstone & Moira Bovill (eds.), *Children and their changing media environment*: *A European comparative study*. Mahwah, NJ: Lawrence Erlbaum Associates, 3–30.

Llanes, Àngels. 2018. The role of language awareness in a study abroad context. In Peter Garrett & Joseph M. Cots (eds.), *The Routledge handbook of language awareness*. New York: Routledge, 275–289.

References

Lo Bianco, Joseph. 2014. Domesticating the foreign: Globalization's effects on the place/s of languages. *The Modern Language Journal* 98(1), 312–325.

Lonergan, Robyn & Cumming, Therese M. 2017. Riding the rapids of classroom-based research. *The Australian Educational Researcher* 44(2), 141–160.

Loschky, Lester. 1994. Comprehensible input and second language acquisition: What is the relationship? *Studies in Second Language Acquisition* 16, 303–323.

Lyrigkou, Christina. 2018. Not to be overlooked: Agency in informal language contact. *Innovation in Language Learning and Teaching* 16(4), 1–16.

Ma, Qing. 2009. *Second language vocabulary acquisition*. Bern: Peter Lang.

Mackey, Alison & Bryfonski, Lara. 2018. Mixed methodology. In Aek Phakiti, Peter de Costa, Luke Plonsky & Sue Starfield (eds.), *The Palgrave handbook of applied linguistics research methodology*. London: Palgrave Macmillan, 103–121.

Mackey, Alison & Gass, Susan M. 2005. *Second language research*: *Methodology and design*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mair, Christian. 2020. English in the German-speaking world: An inevitable presence. In Raymond Hickey (ed.), *English in the German-speaking world*. Cambridge: Cambridge University Press, 13–30.

Malone, Jonathan. 2018. Incidental vocabulary learning in SLA. *Studies in Second Language Acquisition* 40(3), 651–675.

Martinez, Ron & Schmitt, Norbert. 2012. A phrasal expressions list. *Applied Linguistics* 33(3), 299–320.

Mayer, Richard E. 2014. Introduction to multimedia learning. In Richard E. Mayer (ed.), *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press, 1–24.

McLean, Stuart. 2018. Evidence for the adoption of the flemma as an appropriate word counting unit. *Applied Linguistics* 39(6), 823–845.

McLean, Stuart; Kramer, Brandon & Beglar, David. 2015. The creation and validation of a listening vocabulary levels test. *Language Teaching Research* 19(6), 741–760.

Meara, Paul. 1992. *EFL vocabulary tests*. Swansea: University College Swansea: Centre for Applied Language Studies.

Meara, Paul. 1996. The dimensions of lexical competence. In Gillian Brown, Kirsten Malmkjær & John Williams (eds.), *Performance and competence in second language acquisition*. Cambridge: Cambridge University Press, 35–53.

Meara, Paul. 2005. *LLAMA language aptitude tests*: *The manual.* Swansea: lognostics. http://www.lognostics.co.uk/tools/llama/llama_manual.pdf (10 October, 2018).

Meara, Paul. 2009. *Connected words*: *Word associations and second language vocabulary acquisition*. Amsterdam, Philadelphia: John Benjamins Pub. Co.

Meara, Paul. 2010. *EFL vocabulary tests.* Swansea: lognostics. http://www.lognostics.co.uk/vlibrary/index.htm (16 June, 2016).

Meara, Paul. 2012. The bibliometrics of vocabulary acquisition: An exploratory study. *RELC Journal* 43(1), 7–22.

Meara, Paul. 2014. Life before Nation: Bibliometrics and L2 vocabulary studies in 1982. In Gómez González, María de los Ángeles, Ruiz de Mendoza Ibáñez, Francisco José, Francisco Gonzálvez-García & Angela Downing (eds.), *The Functional perspective on language and discourse*: *Applications and implications*. Amsterdam: John Benjamins, 111–129.

Meara, Paul. 2015a. *V_YesNo: a Yes/No vocabulary test for English (v1.01).* Swansea: lognostics. http://www.lognostics.co.uk/tools/V_YesNo/V_YesNo.htm (18 November, 2015).

Meara, Paul. 2015b. Vocabulary research in 1983: A bibliometric analysis. *Linguistics Beyond And Within* 1, 187–198.

Meara, Paul. 2016. Two steps backwards: A bibliometric analysis of L2 vocabulary research in 1984. *Linguistics Beyond And Within* 2, 139–152.

Meara, Paul. 2017. A new beginning? A bibliometric analysis of L2 vocabulary research in 1985. *Linguistics Beyond And Within* 3, 136–154.

Meara, Paul & Buxton, Barbara. 1987. An alternative to multiple choice vocabulary tests. *Language Testing* 4(2), 142–154.

Meara, Paul & Fitzpatrick, Tess. 2000. Lex30: An improved method of assessing productive vocabulary in an L2. *System* 28(1), 19–30.

Meara, Paul & Jones, Glyn. 1988. Vocabulary size as a placement indicator. In Pamela Grunwell (ed.), *Applied linguistcs in society*. London: CILT, 80–87.

Meara, Paul & Jones, Glyn. 1990. *Eurocentres vocabulary size tests*. Zurich: Eurocentres Learning Service.

Meara, Paul & Milton, James. 2003. *X-Lex*: *The Swansea levels test*. Newbury: Express Publishing.

Meara, Paul & Miralpeix, Imma. 2017. *Tools for researching vocabulary*. Bristol: Multilingual Matters.

Meara, Paul & Olmos Acoy, Juan C. 2010. Words as species: An alternative approach to estimating productive vocabulary size. *Reading in a Foreign Language* 22(1), 222–236.

Meara, Paul & Wolter, Brent. 2004. V_links: Beyond vocabulary depth. *Angles on the English-speaking World* 4, 85–96.

Media Consulting Group/EACEA. 2009. *Study on the use of subtitling*: *The potential of subtitling to encourage foreign language learning and improve the mastery of foreign languages.* Paris. https://publications.europa.eu/en/publication-detail/-/publication/e4d5cbf4-a839-4a8a-81d0-7b19a22cc5ce/language-en (17 July, 2018).

Miglbauer, Marlene. 2017. *Students' extramural English as a resource for fostering language skills and digital competencies in tertiary language education.* Krems: Danube University Krems, MA thesis.

Milton, James. 2008. Vocabulary uptake from informal learning tasks. *Language Learning Journal* 36(2), 227–237.

Milton, James. 2009. *Measuring second language vocabulary acquisition*. Bristol: Multilingual Matters.

Milton, James & Alexiou, Thomaï. 2009. Vocabulary size and the Common European Framework of Reference for languages. In Brian Richards, Michael H. Daller, David D. Malvern, Paul Meara, James Milton & Jeanine Treffers-Daller (eds.), *Vocabulary studies in first and second language acquisition*: *The interface between theory and application*. Houndsmill, Basingstoke: Palgrave Macmillan, 194–211.

Milton, James; Alexiou, Thomaï & Mattheoudakis, Marina. 2014. Knowledge of spoken form. In James Milton & Tess Fitzpatrick (eds.), *Dimensions of vocabulary knowledge*. Houndsmill, Basingstoke: Palgrave Macmillan, 13–29.

Milton, James & Fitzpatrick, Tess. 2014. Introduction: Deconstructing vocabulary knowledge. In James Milton & Tess Fitzpatrick (eds.), *Dimensions of vocabulary knowledge*. Houndsmill, Basingstoke: Palgrave Macmillan, 1–12.

Milton, James & Hopkins, Nicola. 2005. *Aural lex*. Swansea: University of Wales Swansea.

Milton, James & Hopkins, Nicola. 2006. Comparing phonological and orthographic vocabulary size: Do vocabulary tests underestimate the knowledge of some learners. *The Canadian Modern Language Review* 63(1), 127–147.

Milton, James & Meara, Paul. 1998. Are the British really bad at learning foreign languages? *Language Learning Journal* 18(1), 68–76.

References

Milton, James & Treffers-Daller, Jeanine. 2013. Vocabulary size revisited: The link between vocabulary size and academic achievement. *Applied Linguistics Review* 4(1), 151–172.

Milton, James; Wade, Jo & Hopkins, Nicola. 2010. Aural word recognition and oral competence in English as a foreign language. In Rubén Chacón-Beltrán, Christián Abello-Contesse & María d. M. Torreblanca-López (eds.), *Insights into non-native vocabulary teaching and learning*. Bristol: Multilingual Matters, 83–98.

Miralpeix, Imma & Muñoz, Carmen. 2018. Receptive vocabulary size and its relationship to EFL language skills. *International Review of Applied Linguistics in Language Teaching* 56(1), 1–24.

Mirmán Flores, Ana & García Jiménez, Eduardo. 2018. The influence of family environment on exposure to English among Spanish secondary school students. *Estudios sobre Educación* 34, 283–306.

Mitchell, Rosamund; Myles, Florence & Marsden, Emma. 2013. *Second language learning theories*, 3rd edn. Abingdon, Oxon: Routledge.

Mizumoto, Atsushi & Plonsky, Luke. 2016. R as a lingua franca: Advantages of using R for quantitative research in applied linguistics. *Applied Linguistics* 37(2), 284–291.

Mizumoto, Atsushi & Takeuchi, Osamu. 2012. Adaptation and validation of self-regulating capacity in vocabulary learning scale. *Applied Linguistics* 33(1), 83–91.

Mochida, Akira & Harrington, Michael. 2006. The Yes/No test as a measure of receptive vocabulary knowledge. *Language Testing* 23(1), 73–98.

Mohamed, Ayman A. 2018. Exposure frequency in L2 reading: An eye-movement perspective on incidental vocabulary learning. *Studies in Second Language Acquisition* 40(2), 269–293.

Mohsen, Mohammed A. 2016. The use of computer-based simulation to aid comprehension and incidental vocabulary learning. *Journal of Educational Computing Research* 54(6), 863–884.

Moir, Jo & Nation, I.S.P. 2002. Learners' use of strategies for effective vocabulary learning. *Prospect* 17(1), 15–35.

Monteleone, Sarah & Puccio, Laura. 2017. *From safe harbour to privacy shield*: *Advances and shortcomings of the new EU-US data transfer rules:* European Parliament. http://www.europarl.europa.eu/RegData/etudes/IDAN/2017/595892/EPRS_IDA(2017)595892 _EN.pdf (20 March, 2019).

Montero Perez, Maribel; Peters, Elke; Clarebout, Geraldine & Desmet, Piet. 2014. Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology* 18(1), 118–141.

Montero Perez, Maribel; Peters, Elke & Desmet, Piet. 2017. Vocabulary learning through viewing video: The effect of two enhancement techniques. *Computer Assisted Language Learning* 31(1-2), 1–26.

Montero Perez, Maribel; Van Den Noortgate, Wim & Desmet, Piet. 2013. Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System* 41(3), 720–739.

Moreno Espinosa, Soraya. 2010. Young learners' L2 word association responses in two different learning contexts. In Jennifer A. Sandlin, Brian D. Schultz & Jake Burdick (eds.), *Handbook of public pedagogy*: *Education and learning beyond schooling*. New York: Routledge, 241–256.

Morgan, David L. 2001. Focus group interviewing. In Jaber F. Gubrium & Holstein James A. (eds.)*, Handbook of interview research*: *Context and method*. Thousand Oaks: SAGE, 141–159.

Morse, Janice M. 1991. Approaches to qualitative-quantitative methodological triangulation. *Nursing Research* 40(2), 120-123.

Morse, Janice M. 2003. Principles of mixed methods and multimethod research design. In Abbas Tashakkori & Charles Teddlie (eds.), *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: SAGE, 189–208.

Mueller, Charles W. & Parcel, Toby L. 1981. Measures of socioeconomic status: Alternatives and recommendations. *Child Development* 52, 13–30.

Mukundan, Jayakaran; Khojasteh, Laleh & Pearson, Nigel. 2009. Exploring the language learning materials used out-of-class by Malaysian TESL students and students of TBFL. *Indonesian Journal of English Language Teaching* 5(1), 40–56.

Mummendey, Hans D. & Grau, Ina. 2014. *Die Fragebogen-Methode* [The questionnaire method], 6th edn. Göttingen: Hogrefe.

Munoz, Carmen. 2008. Symmetries and asymmetries of age effects in naturalistic and instructed L2 Learning. *Applied Linguistics* 29(4), 578–596.

Muñoz, Carmen. 2011. Input and long-term effects of starting age in foreign language learning. *IRAL - International Review of Applied Linguistics in Language Teaching* 49(2), 113–133.

Muñoz, Carmen. 2012. The significance of intensive exposure as a turning point in learners' histories. In Carmen Muñoz (ed.), *Intensive exposure experiences in second language learning*. Bristol: Multilingual Matters, 141–160.

Muñoz, Carmen. 2013. Explicit learning in second language acquisition. In Carol A. Chapelle (ed.), *The encyclopedia of applied linguistics*. Oxford, UK: Blackwell, n.d.

Muñoz, Carmen. 2014. Exploring young learners' foreign language learning awareness. *Language Awareness* 23(1-2), 24–40.

Murray, Garold. 2011. Older language learners, social learning spaces and community. In Phil Benson & Hayo Reinders (eds.), *Beyond the language classroom*. Houndsmill, Basingstoke: Palgrave Macmillan, 132–145.

Nagel, Tanja; Schad, Anke; Semmler, Barbara & Wimmer, Michael. 2012. Austria. In Guus Extra & Kutlay Yağmur (eds.), *Language rich Europe*: *Trends in policies and practices for multilingualism in Europe*. Cambridge: Cambridge University Press, 83–90.

Nagy, William E.; Anderson, Richard C.; Schommer, Marlene; Scott, Judith A. & Stallman, Anne C. 1989. Morphological families in the internal lexicon. *Reading Research Quarterly* 24(3), 262–282.

Nassaji, Hossein. 2003. L2 vocabulary learning from context: Strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *TESOL Quarterly* 37(4), 645–670.

Nassaji, Hossein & Hu, Hsueh-chao M. 2012. The relationship between task-induced involvement load and learning new words from context. *International Review of Applied Linguistics in Language Teaching* 50(1), 69–86.

Nation, I.S.P. 1983. Testing and teaching vocabulary. *Testing and teaching vocabulary* 5, 12–25.

Nation, I.S.P. 1984. *Vocabulary lists*. Wellington: Victoria University of Wellington, English Language Institute.

Nation, I.S.P. 1990. *Teaching and learning vocabulary*. Boston, MA: Heinle & Heinle.

Nation, I.S.P. 2001. *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, I.S.P. 2006. How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review* 63(1), 59–82.

Nation, I.S.P. 2007. Fundamental issues in modelling and assessing vocabulary knowledge. In Helmut Daller, James Milton & Jeanine Treffers-Daller (eds.), *Modelling and assessing vocabulary knowledge*. Cambridge: Cambridge University Press, 35–43.

References

Nation, I.S.P. 2012a. *The BNC/COCA word family lists.* Wellington. https://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/Information-on-the-BNC_COCA-word-family-lists.pdf (9 March, 2017).

Nation, I.S.P. 2012b. *The vocabulary size test.* Wellington. www.victoria.ac.nz/lals/about/staff/publications/paul-nation/Vocabulary-Size-Test-information-and-specifications.pdf (20 October, 2015).

Nation, I.S.P. 2013. *Learning vocabulary in another language*, 2nd edn. Cambridge: Cambridge University Press.

Nation, I.S.P. & Beglar, David. 2007. A vocabulary size test. *The Language Teacher* 31(7), 9–13.

Nation, I.S.P. & Webb, Stuart. 2011. *Researching and analyzing vocabulary*. Boston, MA: Heinle Cengage Learning.

Neuman, Susan B. & Koskinen, Patricia. 1992. Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly* 27(1), 94–106.

Nightingale, Richard. 2016. *The effect of out-of-school media contact on language attitudes in multilingual adolescents*: *A complex psycho-sociolinguistic system.* Castelló de la Plana: Jaume I University, PhD thesis.

Noughabi, Mostafa A. 2017. The effect of meaning-focused listening input on Iranian intermediate EFL learners' productive vocabulary size. *Journal of Education and Practice* 8(5), 141–149.

Nunan, David & Richards, Jack C. (eds.). 2015a. *Language learning beyond the classroom*. New York: Routledge.

Nunan, David & Richards, Jack C. 2015b. Preface. In David Nunan & Jack C. Richards (eds.), *Language learning beyond the classroom*. New York: Routledge, xi–xvi.

Nurmukhamedov, Ulugbek. 2017. Lexical coverage of TED talks: Implications for vocabulary instruction. *TESOL Journal* 8(4), 768–790.

Nurweni, Ari & Read, John. 1999. The English vocabulary knowledge of Indonesian university students. *English for Specific Purposes* 18(2), 161–175.

Nusche, Deborah; Radinger, Thomas; Busemeyer, Marius R. & Theisens, Henno. 2016. *OECD reviews of school resources*: *Austria 2016.* Paris: OECD Publishing. https://www.oecd-ilibrary.org/education/oecd-reviews-of-school-resources-austria-2016_9789264256729-en (10 January, 2018).

O'Loughlin, Richard. 2012. Tuning in to vocabulary frequency in coursebooks. *RELC Journal* 43(2), 255–269.

O'Cathain, Alicia. 2010. Assessing the quality of mixed methods research: Toward a comprehensive framework. In Abbas Tashakkori & Charles Teddlie (eds.), *Handbook of mixed methods in social and behavioral research*, 2nd edn. Thousand Oaks, CA: SAGE, 531–555.

O'Cathain, Alicia; Murphy, Elizabeth & Nicholl, Jon. 2008. The quality of mixed methods studies in health services research. *Journal of health services research & policy* 13(2), 92–98.

OECD. 2014. *PISA 2012 technical report:* OECD Publishing. https://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf (15 February, 2016).

OECD. 2016. *PISA 2015 results*: *Excellence and equity in education (Volume I).* Paris: OECD Publishing.

OECD/Bifie. n.d. *Internationaler und nationaler Schülerfragebogen PISA 2012 Österreich*: *Core components* [International and national student questionnaire PISA 2012 Austria]. https://www.bifie.at/material/internationale-studien/pisa/pisa-2012/ (31 March, 2016).

References

Olin-Scheller, Christina & Wikström, Patrik. 2010. Literary prosumers: Young people's reading and writing in a new media landscape. *Education Inquiry* 1(1), 41–56.

Olsson, Eva. 2012. *"Everything I read on the internet is in English"*: *On the impact of extramural English on Swedish 16-year-old pupils' writing proficiency*. ROSA 15. Göteborg. http://www.kultur.gu.se/digitalAssets/1325/1325748_olsson-eva-lic.pdf (7 July, 2014).

Olsson, Eva. 2016. *On the impact of extramural English and CLIL on productive vocabulary.* Göteborg: University of Gothenburg, PhD thesis.

Olsson, Eva & Sylvén, Liss K. 2015. Extramural English and academic vocabulary: A longitudinal study of CLIL and non-CLIL students in Sweden. *Apples – Journal of Applied Language Studies* 9(2), 77–103.

Onwuegbuzie, Anthony J. & Johnson, R. B. 2006. The validity issue in mixed research. *Research in the Schools* 13(1), 48–63.

Onwuegbuzie, Anthony J.; Johnson, R. B. & Collins, Kathleen M. T. 2011. Assessing legitimation in mixed research: A new framework. *Quality & Quantity* 45(6), 1253–1271.

Ortega, Lourdes. 2011. Second language acquisition. In James Simpson (ed.), *The Routledge handbook of applied linguistics*, 1st edn. Milton Park, Abingdon: Routledge, 171–184.

Oscarson, Mats. 2014. Self-assessment in the classroom. In Anthony J. Kunnan (ed.), *The companion to language assessment*: *Volume II. Approaches and development*. Malden, MA: Wiley-Blackwell, 712–729.

Österreichischer Gehörlosenbund [Austrian Federation for the Deaf]. n.d. *FAQ 1 - Allgemeines zur Gebärdensprache* [FAQ 1- General information on sign language]. http://www.oeglb.at/gebaerdensprache/faq-1-allgemeines-zur-gebaerdensprache/ (9 May, 2018).

Österreichischer Rundfunk Austrian Broadcasting Corporation. n.d., Household equipment and reception. *Haushaltsausstattung und Empfang* [Household equipment and reception]. http://medienforschung.orf.at/medienforschung/fernsehen/technikhaushalt/index.html (12 July, 2017).

Österreichischer Rundfunk Austrian Broadcasting Corporation. n.d. *Information in English*. http://der.orf.at/unternehmen/orf-english100.html (17 May, 2018).

Österreichischer Rundfunk Austrian Broadcasting Corporation. n.d., Information on the programme of FM4. *Informationen zum FM4 Programm* [Information on the programme of FM4]. http://fm4v3.orf.at/radio/stories/faq.html#newssprachen (17 May, 2018).

O'Sullivan, Barry & Weir, Cyril. 2011. Test development and validation. In Barry O'Sullivan (ed.), *Language testing*: *Theories and practices*. Basingstoke: Palgrave Macmillan, 13–32.

Overwien, Bernd. 2005. Stichwort: Informelles Lernen [Keyword: Informal learning]. *Zeitschrift für Erziehungswissenschaft* 8(3), 339–355.

Overwien, Bernd. 2016. Informelles Lernen - Ein historischer Abriss [Informal learning - a historical outline]. In Marius Harring, Matthias D. Witte & Timo Burger (eds.), *Handbuch informelles Lernen*: *Interdisziplinäre und internationale Perspektiven.* [Handbook informal learning: Interdisciplinary and international perspectives]. Weinheim: Beltz Juventa, 41–51.

Ozturk, Meral. 2015. Vocabulary growth of the advanced EFL learner. *The Language Learning Journal* 43(1), 94–109.

Ozturk, Meral. 2016. Second language vocabulary growth at advanced level. *The Language Learning Journal* 44(1), 6–16.

Paivio, Allan. 1986. *Mental representations*: *A dual coding approach*. Oxford: Oxford University Press.

Paivio, Allan. 2007. *Mind and its evolution*: *A dual coding theoretical approach*. Mahwah, NJ: Lawrence Erlbaum Associates.

References

Paribakht, T. S. 2005. The influence of first language lexicalization on second language lexical inferencing: A study of Farsi-speaking learners of English as a foreign language. *Language Learning* 55(4), 701–748.

Paribakht, T. S. & Wesche, Marjorie. 1993. Reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal* 11(1), 9–29.

Paribakht, T. S. & Wesche, Marjorie. 1997. Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In James Coady & Thomas N. Huckin (eds.), *Second language vocabulary acquisition*: *A rationale for pedagogy*. Cambridge: Cambridge University Press, 174–200.

Parlamentsdirektion [Parliamentary Administration]. n.d. *Die Rechte der Volksgruppen* [The rights of the ethnic minorities]. https://www.parlament.gv.at/PERK/VERF/VOLK/ (9 May, 2018).

Pavakanun, Ubolwanna & d'Ydewalle, Géry. 1992. Watching foreign television programs and language learning. In Frits L. Engel, Don G. Bouwhuis, Tom Bösser & Géry d'Ydewalle (eds.), *Cognitive modelling and interactive environments in language learning*. Berlin: Springer, 193–198.

Pavičić Takač, Višnja. 2008. *Vocabulary learning strategies and foreign language acquisition*. Clevedon: Multilingual Matters.

Pearson, Nigel. 2004. *The idiosyncrasies of out-of-class language learning: A study of mainland Chinese students studying English at tertiary level in New Zealand*. Proceedings of the Independent Learning Conference 2003: Independent Learning Association. http://www.independentlearning.org/proceedings-melbourne-2003.html (5 October, 2014).

Pellicer-Sánchez, Ana. 2016. Incidental L2 vocabulary acquisition from and while reading: An eye-tracking study. *Studies in Second Language Acquisition* 38(1), 97–130.

Pellicer-Sánchez, Ana. 2017. Learning L2 collocations incidentally from reading. *Language Teaching Research* 21(3), 381–402.

Pellicer-Sánchez, Ana & Schmitt, Norbert. 2010. Incidental vocabulary acquisition from an authentic novel: Do Things Fall Apart? *Reading in a Foreign Language* 22(1), 31–55.

Pellicer-Sánchez, Ana & Schmitt, Norbert. 2012. Scoring yes-no vocabulary tests: Reaction time vs. nonword approaches. *Language Testing* 29(4), 489–509.

Pellicer-Sánchez, Ana & Siyanova-Chanturia, Anna. 2018. Eye movements in vocabulary research. *ITL - International Journal of Applied Linguistics* 169(1), 5–29.

Persson, Liv & Prins, Tineke. 2012. Learning English inside and outside the classroom. In Nel De Jong, Kasper Juffermans, Merel Keijzer & Laurent Rasier (eds.), *Papers of the Anéla 2012 applied linguistics conference*. Delft: Uitgeverij Eburon, 3–13.

Peters, Elke. 2018. The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *ITL - International Journal of Applied Linguistics* 169(1), 142–168.

Peters, Elke; Heynen, Eva & Puimège, Eva. 2016. Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System* 63, 134–148.

Peters, Elke; Noreillie, Ann-Sophie; Heylen, Kris; Bulté, Bram & Desmet, Piet. 2019. The impact of instruction and out-of-school exposure to foreign language input on learners' vocabulary knowledge in two languages. *Language Learning* 69(3), 747–782.

Peters, Elke; Velghe, Tom & Van Rompaey, Tinne. 2015. A post-entry English and French vocabulary size for Flemish learners (Paper presented at EALTA 2015. Copenhagen, 29 May 2015).

Peters, Elke & Webb, Stuart. 2018. Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition* 40, 551–577.

References

Petrescu, Maria C.; Helms-Park, Rena & Dronjic, Vedran. 2017. The impact of frequency and register on cognate facilitation: Comparing Romanian and Vietnamese speakers on the Vocabulary Levels Test. *English for Specific Purposes* 47, 15–25.

Pfenninger, Simone E. & Singleton, David. 2017. *Beyond age effects in instructional L2 learning*: *Revisiting the age factor*. Bristol: Multilingual Matters.

Pham, Giang; Freunberger, Roman & Robitzsch, Alexander. 2014. *Hintergrundvariablen und spezielle Analysen*: *Technische Dokumentation - BIST-Ü Englisch, 8. Schulstufe, 2013* [Background variables and specific analyses: Technical documentation - Educational standards English, grade 8, 2013]. Salzburg: bifie. https://www.bifie.at/material/ueberpruefung-der-bildungsstandards/technische-dokumentation/ (7 September, 2017).

Pickard, Nigel. 1996. Out-of-class language learning strategies. *ELT Journal* 50(2), 150–159.

Pigada, Maria & Schmitt, Norbert. 2006. Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language* 18(1), 1–28.

Pill, Thomas J. H. 2001. *Adult learners' perceptions of out-of-class access to English.* Hong Kong: University of Hong Kong, MA thesis.

Piritidis, Katherina. 2014. *The linguistic landscape of Vienna's Westbahnstraße – A comparison of methods.* Wien: University of Vienna, MA thesis.

Platzer, Hans. 2006. Englischkompetenz unter erstsemestrigen WirtschaftsstudentInnen: Eine empirische Untersuchung [English competence among first-year students of economy. An empirical investigation]. *AAA - Arbeiten aus Anglistik und Amerikanistik* 31(2), 209–236.

Plonsky, Luke. 2013. Study quality in SLA: An assessment of designs, analyses, and reporting practices in quantitative L2 research. *Studies in Second Language Acquisition* 35(4), 655–687.

Plonsky, Luke. 2014. Study quality in quantitative L2 research (1990-2010): A methodological synthesis and call for reform. *The Modern Language Journal* 98(1), 450–470.

Plonsky, Luke. 2015. Statistical power, p values, descriptive statistics, and effect sizes: A "back-to-basics" approach to advancing quantitative methods in L2 research. In Luke Plonsky (ed.), *Advancing quantitative methods in second language research*. New York: Routledge, 23–45.

Plonsky, Luke & Oswald, Frederick L. 2012. How to do a meta-analysis. In Alison Mackey & Susan M. Gass (eds.), *Research methods in second language acquisition*: *A practical guide*. Malden, MA: Blackwell, 275–295.

Plonsky, Luke & Oswald, Frederick L. 2014. How big is "big"? Interpreting effect sizes in L2 research. *Language Learning* 64(4), 878–912.

Podrepschek, Nora. 2016. *Raum für linguistische Diversität in Ottakring? Eine Analyse der sichtbaren Mehrsprachigkeit in Wiens 16. Gemeindebezirk unter besonderer Berücksichtigung der türkischen und bosnisch/kroatisch/serbischen Minderheiten* [Room for linguistic diversity in Ottakring? An analysis of visible multilingualism in Vienna's 16th district with special attention to Turkish and Bosnian/Croatian/Serbian minorities]. Wien: University of Vienna, MA thesis.

Porte, Graeme K. 2010. *Appraising research in second language learning*: *A practical approach to critical analysis of quantitative research*, 2nd edn. Amsterdam: John Benjamins.

Porte, Graeme K. 2012. Introduction. In Graeme K. Porte (ed.), *Replication research in applied linguistics*. Cambridge: Cambridge University Press, 1–18.

Puchta, Herbert; Holzmann, Christian; Stranks, Jeff & Lewis-Jones, Peter. 2013a. *Into English 1*: *Coursebook*. Innsbruck: Helbling Languages.

Puchta, Herbert; Holzmann, Christian; Stranks, Jeff & Lewis-Jones, Peter. 2013b. *Into English 2*: *Coursebook*. Innsbruck: Helbling Languages.

References

Puimège, Eva & Peters, Elke. 2019. Learners' English vocabulary knowledge prior to formal instruction: The role of learner-related and word-related factors. *Language Learning* 69(4), 943–977.

Puka, Llukan. 2011. Kendall's tau. In Miodrag Lovric (ed.), *International encyclopedia of statistical science*. Berlin: Springer, 713–715.

Pulido, Diana. 2003. Modeling the role of second language proficiency and topic familiarity in second language incidental vocabulary acquisition through reading. *Language Learning* 53(2), 233–284.

Pulido, Diana. 2007. The effects of topic familiarity and passage sight vocabulary on L2 lexical inferencing and retention through reading. *Applied Linguistics* 28(1), 66–86.

Purkarthofer, Judith. 2013. Lokal, global und mehrsprachig? Sprachenpolitik und Medien [Local, global and multilingual? Language policy and media]. In Rudolf De Cillia & Eva Vetter (eds.), *Sprachenpolitik in Österreich*: *Bestandsaufnahme 2011*. [Language policy in Austria: Review 2011]. Frankfurt am Main: Peter Lang, 242–256.

Putman, S. M. & Kingsley, Tara. 2009. The Atoms Family: Using podcasts to enhance the development of science vocabulary. *The Reading Teacher* 63(2), 100–108.

Qian, David D. 2005. Demystifying lexical inferencing: The role of aspects of vocabulary knowledge. *TESL Canada Journal* 22, 34–54.

Qualtrics LLC. 2015. *Qualtrics security white paper lite*: *Defining our security processes*. www.qualtrics.com/security-statement (12 May, 2016).

Qualtrics LLC. 2016. *Qualtrics survey software* . Provo, Utah.

R Development Core Team. 2018. *R*: *A language and environment for statistical computing* (3.5.1). Vienna: The R Foundation for Statstical Computing. www.R-project.org/.

Rankin, Yolanda; Gold, Rachel & Gooch, Bruce. 2006. 3D role-playing games as language learning tools. *Europgraphics 2006* 25(3), 1–6.

Ranta, Elina. 2010. English in the real world vs. English at school: Finnish English teachers' and students' views. *International Journal of Applied Linguistics* 20(2), 156–177.

Read, John. 1988. Measuring the vocabulary knowledge of second langauge [sic] learners. *RELC Journal* 19(2), 12–25.

Read, John. 2000. *Assessing vocabulary*. Cambridge: Cambridge University Press.

Read, John. 2004. Research in teaching vocabulary. *Annual Review of Applied Linguistics* 24, 146–161.

Read, John & Chapelle, Carol A. 2001. A framework for second language vocabulary assessment. *Language Testing* 18(1), 1–32.

Rebuschat, Patrick & Williams, John N. 2013. Implicit learning in second language acquisition. In Carol A. Chapelle (ed.), *The encyclopedia of applied linguistics*. Oxford, UK: Blackwell, n.d.

Reinders, Hayo. 2017. Digital games and second language learning. In Steven L. Thorne & Stephen May (eds.), *Encyclopedia of language and education*: *Language, education and technology*, 3rd edn. Cham: Springer, 1–15.

Reinders, Hayo & Benson, Phil. 2017. Research agenda: Language learning beyond the classroom. *Language Teaching* 50(4), 561–578.

Reinders, Hayo & Wattana, Sorada. 2010. Learn English or die: The effects of digital games on interaction and willingness to communicate in a foreign language. *Digital Culture and Education* 3(1), 4–28.

Reinders, Hayo & Wattana, Sorada. 2015. The effects of digital game play on second language interaction. *International Journal of Computer-Assisted Language Learning and Teaching* 5(1), 1–21.

References

Riazi, A. M. 2017. *Mixed methods research in language teaching and learning*. Sheffield: Equinox.

Richards, Jack C. 1976. The role of vocabulary teaching. *TESOL Quarterly* 10(1), 77–89.

Richards, Jack C. 2015. The changing face of language learning: Learning beyond the classroom. *RELC Journal* 46(1), 5–22.

Rieder, Angelika. 2003. Implicit and explicit learning in incidental vocabulary acquisition. *Vienna English Working Papers (VIEWS)* 12(2), 24–39.

Ringl, Barbara. 2014. *Extracurricular use of English among Austrian adolescents.* Wien: University of Vienna, MA thesis.

Rodgers, Michael P. H. 2013. *English language learning through viewing television*: *An investigation of comprehension, incidental vocabulary acquisition, lexical coverage, attitudes, and captions.* Wellington: Victoria University of Wellington, PhD thesis.

Rodgers, Michael P. H. 2018. The images in television programs and the potential for learning unknown words: The relationship between on-screen imagery and vocabulary. *ITL - International Journal of Applied Linguistics* 169(1), 191–211.

Rodgers, Michael P. H.; Heidt, Julian & Wood, David. 2019. Game on for comprehensible input and vocabulary learning: The lexical demands of video games (Paper presented at Vocab@Leuven 2019. Leuven, 1 July 2019).

Rodgers, Michael P. H. & Webb, Stuart. 2017. The effects of captions on EFL learners' comprehension of English-language television programs. *CALICO Journal* 34(1), 20–38.

Rogers, Alan. 2008. Informal learning and literacy. In Brian V. Street & Nancy H. Hornberger (eds.), *Encyclopedia of language and education*: *Volume 2. Literacy,* 2nd edn. New York: Springer, 133–144.

Rogers, Vivienne E.; Meara, Paul; Aspinall, Rachel; Fallon, Louise; Goss, Thomas; Keey, Emily & Thomas, Rosa. 2016. Testing aptitude: Investigating Meara's (2005) LLAMA tests. *EUROSLA Yearbook* 16(1), 179–210.

Rogers, Vivienne E.; Meara, Paul; Barnett-Legh, Thomas; Curry, Clare & Davie, Emma. 2017. Examining the LLAMA aptitude tests. *Journal of the European Second Language Association* 1(1), 49–60.

Rohs, Matthias. 2010. Zur Neudimensionierung des Lernortes [On the redimensioning of the learning space]. *REPORT Zeitschrift für Weiterbildungsforschung*(2), 34–45.

Rohs, Matthias. 2016a. Genese informellen Lernens [The genesis of informal learning]. In Matthias Rohs (ed.), *Handbuch informelles Lernen.* [Handbook informal learning]. Wiesbaden: Springer, 3–38.

Rohs, Matthias (ed.). 2016b. *Handbuch informelles Lernen* [Handbook informal learning]. Wiesbaden: Springer.

Rose, Heath; Briggs, Jessica G.; Boggs, Jill A.; Sergio, Lia & Ivanova-Slavianskaia, Natalia. 2018. A systematic review of language learner strategy research in the face of self-regulation. *System* 72, 151–163.

Rose, Heath & McKinley, Jim. 2017. Realities of doing research in applied linguistics. In Jim McKinley & Heath Rose (eds.), *Doing research in applied linguistics*: *Realities, dilemmas, and solutions*. Abingdon, Oxon: Routledge, 3–14.

Rossiter, Marian J. 2001. The challenges of classroom-based SLA research. *Applied Language Learning* 12(1), 31–44.

RStudio Team. 2018. *RStudio*: *Integrated Development for R* (1.2.1335). Boston, MA: RStudio, Inc. www.rstudio.com/.

Rubin, Joan & Thompson, Irene. 1982. *How to be a more successful language learner*. New York: Heinle & Heinle.

Rundfunk und Telekom Regulierungs-GmbH (RTR) [Public Broadcasting and Telecommunications Regulations Ltd]. 2017. *Die Konkurrenz aus dem Netz*: *OTT-Dienste in*

Medien und Telekommunikation [The competition from the net: OTT services in media and telecommunications]. Wien. https://www.rtr.at/de/inf/Konkurrenz_aus_dem_Netz_OTT (15 May, 2018).

Rupérez Micola, Augusto; Bris, Arturo & Banal-Estañol, Albert. 2009. *TV or not TV? Subtitling and English skills.* Social Science Research Network. http://dx.doi.org/10.2139/ssrn.1435259.

Rymarczyk, Jutta. 2015. Museen als außerschulische Lernorte [Museums as learning spaces outside school]. In Eva Burwitz-Melzer, Frank G. Königs & Claudia Riemer (eds.), *Lernen an allen Orten? Die Rolle der Lernorte beim Lehren und Lernen.* [Learning in all places? The role of learning spaces in teaching and learning of languages]. Tübingen: Narr Verlag, 201–210.

Saad, Noor S. M.; Melor, Md Y. & Embi, Mohamed A. 2013. The intersection between out-of-class language learning strategies and in-class activities. *Advances in Language and Literary Studies* 4(2), 132–140.

Saferinternet.at. 2018. *Jugend-Internet-Monitor 2018* [Youth internet monitor 2018]. Wien. https://www.saferinternet.at/jugendinternetmonitor/ (6 June, 2018).

Salchegger, Silvia; Wallner-Paschon, Christina; Schmich, Juliane & Höller, Iris. 2016. Kompetenzentwicklung im Kontext individueller, schulischer und familiärer Faktoren [The development of competences in the context of individual, school and family factors]. In Birgit Suchań & Simone Breit (eds.), *PISA 2015*: *Grundkompetenzen am Ende der Pflichtschulzeit im internationalen Vergleich.* [PISA 2015: An international comparison of basic competences at the end of compulsory schooling]. Graz: Leykam, 77–100.

Saldaña, Johnny. 2016. *The coding manual for qualitative researchers*, 3rd edn. Thousand Oaks: SAGE.

Sarkeshikian, Amir H.; Tabatabaee, Abdol-Majid & Doaee, Maryam T. 2018. Unidimensionality and construct validity of the self-regulating capacity in vocabulary learning (SRCvoc) in Iranian EFL context: Item-level responses versus item parcels. *Psychology of Language and Communication* 22(1), 21–38.

Sauro, Shannon. 2017. Online fan practices and CALL. *CALICO Journal* 34(2), 131–146.

Schlick, Maria. 2002. The English of shop signs in Europe: A case study of foreign and especially English influence on the language of shop signs and shop windows in three European cities. *English Today* 18(2), 3–7.

Schlick, Maria. 2003. The English of shop signs in Europe: A case study of the language in the store fronts of eight town centres. *English Today* 19(1), 3–17.

Schmidt, Richard. 1994. Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review* 11, 11–26.

Schmitt, Norbert. 1997. Vocabulary learning strategies. In Norbert Schmitt & Michael McCarthy (eds.), *Vocabulary*: *Description, acquisition and pedagogy*. Cambridge: Cambridge University Press, 199–227.

Schmitt, Norbert. 1998. Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning* 48(2), 281–317.

Schmitt, Norbert. 2000. *Vocabulary in language teaching*. Cambridge: Cambridge University Press.

Schmitt, Norbert (ed.). 2004. *Formulaic sequences*: *Acquisition, processing and use*. Amsterdam: John Benjamins.

Schmitt, Norbert. 2008. Review article: Instructed second language vocabulary learning. *Language Teaching Research* 12(3), 329–363.

Schmitt, Norbert. 2010. *Researching vocabulary*: *A vocabulary research manual*. Basingstoke: Palgrave Macmillan.

References

Schmitt, Norbert. 2014. Size and depth of vocabulary knowledge: What the research shows. *Language Learning* 64(4), 913–951.

Schmitt, Norbert. 2019. Understanding vocabulary acquisition, instruction, and assessment: A research agenda. *Language Teaching* 52(2), 261–274.

Schmitt, Norbert; Cobb, Tom; Horst, Marlise & Schmitt, Diane. 2017. How much vocabulary is needed to use English? Replication of van Zeeland & Schmitt (2012), Nation (2006) and Cobb (2007). *Language Teaching* 50(2), 212–226.

Schmitt, Norbert; Jiang, Xiangying & Grabe, William. 2011. The percentage of words known in a text and reading comprehension. *The Modern Language Journal* 95(1), 26–43.

Schmitt, Norbert & Redwood, Stephen. 2011. Learner knowledge of phrasal verbs: A corpus-informed study. In Fanny Meunier, Sylvie de Cock, Gaëtanelle Gilquin & Magali Paquot (eds.), *A taste for corpora*: *In honour of Sylviane Granger*. Amsterdam: John Benjamins Publishing Company, 137–207.

Schmitt, Norbert & Schmitt, Diane. 2014. A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching* 47(4), 484–503.

Schmitt, Norbert; Schmitt, Diane & Clapham, Caroline. 2001. Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing* 18(1), 55–88.

Schmitt, Norbert & Zimmermann, Cheryl B. 2002. Derivative word forms: What do learners know? *TESOL Quarterly* 36(2), 145–171.

Schneider, Grant; Chicken, Eric & Becvarik, Rachel. 2018. *NSM3: Functions and datasets to accompany Hollander, Wolfe, and Chicken - Nonparametric Statistical Methods, 3rd edition* (R package version 1.12.).

Schöpper-Grabe, Sigrid. 2009. Betrieblicher Fremdsprachenbedarf im deutschsprachigen Raum [The need for foreign languages in business in the German-speaking area]. *Sociolinguistica* 23(1), 150–162.

Schreier, Margrit. 2012. *Qualitative content analysis in practice*. London: SAGE.

Schreier, Margrit. 2014. Qualitative content analysis. In Uwe Flick (ed.), *The SAGE handbook of qualitative data analysis*. Los Angeles: SAGE, 170–183.

Schreiner, Claudia & Breit, Simone. 2014. *Standardüberprüfung 2013 Englisch, 8. Schulstufe*: *Bundesergebnisbericht* [Educational standards 2013 English, Grade 8: Results for the federal state]. Salzburg: Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens [Federal Institute of Educational Research, Innovation and Development of the Austrian School Sector]. https://www.bifie.at/ergebnisberichte/ (14 May, 2018).

Schwarz, Marlene. 2013. Learning with Lady GaGa & Co: Incidental EFL vocabulary acquisition from pop songs. *Vienna English Working Papers (VIEWS)* 22, 17–48.

Schwarz, Marlene. 2016. Beyond the walls – Vocabulary learning from extramural English. In Monika Boniecki (ed.), *ÖGSD Tagungsberichte Vol. 1: 8. Nachwuchstagung*: *Sprachendidaktik: Der wissenschaftliche Nachwuchs im Dialog. (Proceedings of the 8th ÖGSD Young Researchers' Conference)* (1). Graz: ÖGSD, 58–61.

Sefton-Green, Julian. 2004. *Literature review in informal learning with technology outside school*. Bristol: Futurelab.

Sefton-Green, Julian. 2013. *Learning at not-school*: *A review of study, theory, and advocacy for education in non-formal settings*. Cambridge, MA: MIT Press.

Seidlhofer, Barbara. 2001. Closing a conceptual gap: The case for a description of English as a lingua franca. *International Journal of Applied Linguistics* 11(2), 133–158.

Seidlhofer, Barbara. 2011. *Understanding English as a lingua franca*. Oxford: Oxford University Press.

References

Shen, Li-Bi; Tseng, Ching-Ya; Kuo, Shu-Wei; Su, Ying-Ju & Chen, Ming-Yuan. 2005. A preliminary study of college students' out-of-class English learning activities. *CHIA-NAN ANNUAL BULLETIN* 31, 464–475.

Shillaw, John. 1995. Using a word list as a focus for vocabulary learning. *The Language Teacher* 19(2), 58–59.

Silbereisen, Rainer K. 2003. Contextual constraints on adolescents' leisure. *New directions for child and adolescent development* 99, 95–101.

Siyanova-Chanturia, Anna & Pellicer-Sánchez, Ana (eds.). 2019. *Understanding formulaic language*: *A second language acquisition perspective*. New York: Routledge.

Smidt, Esther & Hegelheimer, Volker. 2004. Effects of online academic lectures on ESL listening comprehension, incidental vocabulary acquisition, and strategy use. *Computer Assisted Language Learning* 17(5), 517–556.

Smit, Ute. 2004. Language policies in Viennese schools: The roles English plays. *Vienna English Working Papers (VIEWS)* 13(2), 69–87.

Smit, Ute & Finker, Thomas. 2018. CLIL in Austrian technical colleges ('HTL'): An analysis of classroom practices based on systematic video-based lesson observations. In Monika Dannerer & Peter Mauser (eds.), *Formen der Mehrsprachigkeit*: *Sprachen und Varietäten in sekundären und tertiären Bildungskontexten.* [Forms of multilingualim: Languages and varieties in secondary and tertiary educational contexts]. Tübingen: Stauffenburg Verlag, 229–246.

Smit, Ute & Schwarz, Marlene. 2020. English in Austria: Policies and practices. In Raymond Hickey (ed.), *English in the German-speaking world.* Cambridge: Cambridge University Press, 294–314.

Sockett, Geoffrey. 2011. From the cultural hegemony of English to online informal learning: Cluster frequency as an indicator of relevance in authentic documents. *ASp - la revue du GERAS* 60, 5–20.

Sockett, Geoffrey. 2013. Understanding the online informal learning of English as a complex dynamic system: An emic approach. *ReCALL* 25(1), 48–62.

Sockett, Geoffrey. 2014. *The online informal learning of English*. New York: Palgrave Macmillan.

Sockett, Geoffrey & Toffoli, Denyze. 2012. Beyond learner autonomy: A dynamic systems view of the informal learning of English in virtual online communities. *ReCALL* 24(2), 138–151.

Sonbul, Suhad & Schmitt, Norbert. 2013. Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning* 63(1), 121–159.

Soukup, Barbara. 2016. English in the linguistic landscape of Vienna, Austria (ELLViA): Outline, rationale, and methodology of a large-scale empirical project on language choice on public signs from the perspective of sign-readers. *Vienna English Working Papers (VIEWS)* 25, 1–24.

Soukup, Barbara & Moosmüller, Sylvia. 2011. Standard language in Austria. In Tore Kristiansen & Nikolas Coupland (eds.), *Standard languages and language standards in a changing Europe*. Oslo: Novus Press, 39–46.

Spratt, Mary; Humphreys, Gillian & Chan, Victoria. 2002. Autonomy and motivation: Which comes first? *Language Teaching Research* 6(3), 245–266.

Stæhr, Lars S. 2008. Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal* 36(2), 139–152.

Stæhr, Lars S. 2009. Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition* 31(4), 577–607.

Statistik Austria [Statistics Austria]. n.d. *STATcube - Statistische Datenbank* STATcube - Statististical Database. Vienna. http://statcube.at/statistik.at/ext/statcube/jsf/terms.xhtml (13 February, 2019).

Statistik Austria [Statistics Austria]. 2009. *Zeitverwendung 2008/09: Ein Überblick über geschlechtsspezifische Unterschiede*: *Endbericht der Bundesanstalt Statistik Österreich an die Bundesministerin für Frauen und Öffentlichen Dienst* [Time use 2008/09: An overview of gender-specific differences. Final report of the Federal Institute Statistics Austria for the Federal Ministery of Women and Public Service]. Vienna. http://statistik.at/web_de/statistiken/menschen_und_gesellschaft/soziales/zeitverwendung/zeitverwendungserhebung/index.html (21 December, 2017).

Statistik Austria [Statistics Austria]. 2010. *Zeitverwendung 2008/09*: *Durchschnittliche Zeitverwendung pro Tag (Montag - Sonntag) aller Personen nach Altersgruppen und zusammengefasster Haupttätigkeit* [Time use 2008/09: Average time use per day (Monday to Sunday) of all persons by age group and summarized main activity]. Vienna. http://statistik.at/web_de/statistiken/menschen_und_gesellschaft/soziales/zeitverwendung/zeitverwendungserhebung/index.html (21 December, 2017).

Statistik Austria [Statistics Austria]. 2018. *Bildung in Zahlen 2016/17*: *Schlüsselindikatoren und Analysen* [Education in figures 2016/17: Key indicators and analyses]. Vienna. http://www.statistik.at/web_de/services/publikationen/5/index.html?includePage=detailedView&sectionName=Bildung%2C+Kultur&pubId=508 (8 May, 2018).

Statistik Austria [Statistics Austria]. 2020. *Bevölkerung* [Population]. Vienna. http://statistik.at/web_de/statistiken/menschen_und_gesellschaft/bevoelkerung/index.html (8 March, 2020).

Statistisches Bundesamt. 2015a. *Wie die Zeit vergeht*: *Ergebnisse zur Zeitverwendung in Deutschland 2012/2013* [How time passes: Results on time use in Germany 2012/13]. Wiesbaden. https://www.destatis.de/DE/Publikationen/Thematisch/EinkommenKonsumLebensbedingungen/Zeitbudgeterhebung/Begleitheft_WieDieZeitVergeht_Presse.html (6 June, 2018).

Statistisches Bundesamt. 2015b. *Zeitverwendungserhebung 2012/2013*: *Aktivitäten in Stunden und Minuten für ausgewählte Personengruppen* [Time use survey 2012/2013: Activities in hours and minutes for selected groups]. Wiesbaden. https://www.destatis.de/DE/Publikationen/Thematisch/EinkommenKonsumLebensbedingungen/Zeitbudgeterhebung/Zeitverwendung.html (6 June, 2018).

Stevens, James P. 2009. *Applied multivariate statistics for the social sciences*, 5th edn. New York, NY: Taylor & Francis.

Stewart, David W. & Shamdasani, Prem N. 2015. *Focus groups*: *Theory and practice*. Thousand Oaks: SAGE.

Stewart, David W.; Shamdasani, Prem N. & Rook, Dennis W. 2009. Group depth interviews: Focus group research. In Debra J. Rog & Leonard Bickman (eds.), *The SAGE handbook of applied social research methods*, 2nd edn. Los Angeles: SAGE, 589–616.

Stubbe, Raymond. 2012. Searching for an acceptable false alarm maximum. *VERB* 1(2), 7–9.

Stubbe, Raymond. 2013. Comparing regression versus correction formula predictions of passive recall test scores from yes-no test results. *Vocabulary Learning and Instruction* 2(1), 39–46.

Stubbe, Raymond & Stewart, Jeffrey. 2012. Optimizing scoring formulas for yes/no vocabulary tests with linear models. *Shiken Research Bulleting* 16(2), 2-7.

Subtirelu, Nicholas. 2013. What (do) learners want (?): A re-examination of the issue of learner preferences regarding the use of 'native' speaker norms in English language teaching. *Language Awareness* 22(3), 270–291.

Suchań, Birgit & Breit, Simone (eds.). 2016. *PISA 2015*: *Grundkompetenzen am Ende der Pflichtschulzeit im internationalen Vergleich* [PISA 2015: An international comparison of basic competences at the end of compulsory schooling]. Graz: Leykam.

References

Suh, Jae-Suk; Wasanasomsithi, Punchalee; Short, Stephen & Majid, Norazman A. 1999. *Out of class learning experiences and students' perceptions of their impact on English conversation skills*. http://eric.ed.gov/?id=ED433715 (27 November, 2015).

Sundqvist, Pia. 2009a. *Extramural English matters*: *Out-of-school English and its impact on Swedish ninth graders' oral proficiency and vocabulary*. Karlstad: Karlstad University, PhD thesis.

Sundqvist, Pia. 2009b. The impact of spare time activities on students' English language skills. In Solveig Granath, Björn Bihl & Elisabeth Wennö (eds.), *Vägar till språk och litteratur*. Karlstad: Karlstad University Press, 63–76.

Sundqvist, Pia. 2013. Categorization of digital games in English language learning studies: Introducing the SSI model. In Linda Bradley & Sylvie Thouësny (eds.), *20 Years of EUROCALL*: *Learning from the past, looking to the future (2013 EUROCALL Conference Proceedings)*. Dublin: Research-publishing.net, 231–237.

Sundqvist, Pia. 2015. About a boy: A gamer and L2 English speaker coming into being by use of self-access. *Studies in Self-Access Learning Journal* 6(4), 352–364.

Sundqvist, Pia. 2019. Commercial-off-the-shelf games in the digital wild and L2 learner vocabulary. *Language Learning & Technology* 23(1), 87–113.

Sundqvist, Pia & Olin-Scheller, Christina. 2013. Classroom vs. extramural English: Teachers dealing with demotivation. *Language and Linguistics Compass* 7(6), 329–338.

Sundqvist, Pia & Sylvén, Liss K. 2012. World of VocCraft: Computer games and Swedish learners' L2 English vocabulary. In Hayo Reinders (ed.), *Digital games in language learning and teaching*. Houndsmill, Basingstoke: Palgrave Macmillan, 189–208.

Sundqvist, Pia & Sylvén, Liss K. 2014. Language-related computer use: Focus on young L2 English learners in Sweden. *ReCALL* 26(1), 3–20.

Sundqvist, Pia & Sylvén, Liss K. 2016. *Extramural English in teaching and learning*: *From theory to practice*. London: Palgrave Macmillan.

Sundqvist, Pia & Wikström, Peter. 2015. Out-of-school digital gameplay and in-school L2 English vocabulary outcomes. *System* 51, 65–76.

Sylvén, Liss K. 2004/2010. *Teaching English or English teaching? On the effects of content and language integrated learning on Swedish learners' incidental vocabulary acquisition*. Göteborg: University of Gothenburg.

Sylvén, Liss K. 2006. How is extramural exposure to English among Swedish school students used in the CLIL classroom? *Vienna English Working Papers (VIEWS)* 15(3), 47–53.

Sylvén, Liss K. & Sundqvist, Pia. 2012a. Gaming as extramural English L2 learning and L2 proficiency among young learners. *ReCALL* 24(3), 302–321.

Sylvén, Liss K. & Sundqvist, Pia. 2012b. Similarities between playing World of Warcraft and CLIL. *APPLES – Journal of Applied Language Studies* 6(2), 113–130.

Szudarski, Paweł & Carter, Ronald. 2016. The role of input flood and input enhancement in EFL learners' acquisition of collocations. *International Journal of Applied Linguistics* 26(2), 245–265.

Tabachnick, Barbara G. & Fidell, Linda S. 2013. *Using multivariate statistics*, 6th edn. Boston, MA: Pearson.

Tahmasbi, Maryam & Farvardin, Mohammad T. 2017. Probing the effects of task types on EFL learners' receptive and productive vocabulary knowledge: The case of involvement load hypothesis. *SAGE Open* 7(3), 1-10.

Tanaka, Mitsuko. 2017. Examining EFL vocabulary learning motivation in a demotivating learning environment. *System* 65, 130–138.

Teddlie, Charles & Tashakkori, Abbas. 2003. Major issues and controversies in the use of mixed methods in the social and behavioral sciences. In Abbas Tashakkori & Charles Teddlie

(eds.), *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: SAGE, 3–50.

Teddlie, Charles & Tashakkori, Abbas. 2006. A general typology of research designs featuring mixed methods. *Research in the Schools* 13(1), 12–28.

Teddlie, Charles & Tashakkori, Abbas. 2009. *Foundations of mixed methods research*: *Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, CA: SAGE.

Teddlie, Charles & Yu, Fen. 2007. Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research* 1(1), 77–100.

Tegge, Friederike. 2017. The lexical coverage of popular songs in English language teaching. *System* 67, 87–98.

Testa, Adrienne C. & Coleman, Lester M. 2006. Accessing research participants in schools: A case study of a UK adolescent sexual health survey. *Health Education Research* 21(4), 518–526.

The Douglas Fir Group. 2016. A transdisciplinary framework for SLA in a multilingual world. *The Modern Language Journal* 100(Supplement 2016), 19–47.

Thole, Werner. 2010. Jugend: Freizeit und Medien [Youth: Leisure and media]. In Heinz-Hermann Krüger & Cathleen Grunert (eds.), *Handbuch Kindheits- und Jugendforschung*. [Handbook childhood and youth research], 2nd edn. Wiesbaden: VS Verlag für Sozialwissenschaften, 727–764.

Thorne, Steven L. 2009. 'Community', semiotic flows, and mediated contribution to activity. *Language Teaching* 42(1), 81–94.

Thorne, Steven L.; Black, Rebecca W. & Sykes, Julie M. 2009. Second language use, socialization, and learning in internet interest communities and online gaming. *The Modern Language Journal* 93, 802–820.

Toffoli, Denyze & Sockett, Geoffrey. 2013. University teachers' perceptions of Online Informal Learning of English (OILE). *Computer Assisted Language Learning*, 1–15.

Toffoli, Denyze & Sockett, Geoffrey. 2014. English language music: Does it help with learning? *Cahiers de l'APLIUT (Association des Professeurs de Langues des IUT)* 32(2), 192–209.

Tran, Mai P. 2017. *Exploring young learners' informal learning of English language*: *A comparative study on the perspectives of 11-13-year-old pupils in Finland and Vietnam*. Kuopio: University of Eastern Finland, MA thesis.

Trinder, Ruth. 2015. *Online informal learning of English*: *How students use technology to supplement classes:* International Conference ICT for Language Learning (4th edn.). http://conference.pixel-online.net/ICT4LL/files/ict4ll/ed0008/FP/2161-ICL1391-FP-ICT4LL8.pdf (20 July, 2017).

Trinder, Ruth. 2017. Informal and deliberate learning with new technologies. *ELT Journal* 71(4), 401–412.

Tritscher-Archan, Sabine. 2008. Fremdsprachenbedarf und -kompetenzen in Österreichs Unternehmen [Foreign language needs and competences in Austrian businesses]. In Sabine Tritscher-Archan (ed.), *Fremdsprachen für die Wirtschaft*: *Analysen, Zahlen, Fakten*. [Foreign languages for the economy: Analyses, figures, facts]. Vienna: Institut für Bildungsforschung der Wirtschaft, 171–190.

Tschirner, Erwin. 2004. Breadth of vocabulary and advanced English study: An empirical investigation. *Electronic Journal of Foreign Language Teaching* 1(1), 27–39.

Tseng, Wen-Ta; Dörnyei, Zoltán & Schmitt, Norbert. 2006. A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics* 27(1), 78–102.

References

Tseng, Wen-Ta & Schmitt, Norbert. 2008. Toward a model of motivated vocabulary learning: A structural equation modeling approach. *Language Learning* 58(2), 357–400.

Turgut, Yıldız & İrgin, Pelin. 2009. Young learners' language learning via computer games. *Procedia - Social and Behavioral Sciences* 1(1), 760–764.

Uemura, Toshihiko & Ishikawa, Shin'ichiro. 2004. JACET 8000 and Asia TEFL vocabulary initiative. *Journal of Asia TEFL* 1(1), 333–347.

Uitto, Anna; Juuti, Kalle; Lavonen, Jari & Meisalo, Veijo. 2006. Students' interest in biology and their out-of-school experiences. *Journal of Biological Education* 40(3), 124–129.

UNESCO Institute for Statistics. 2012. *International Standard Classification of Education (ISCED) 2011.* Montreal: UNESCO. http://uis.unesco.org/en/topic/international-standard-classification-education-isced (9 October, 2017).

Ushioda, Ema. 2013. Motivation and ELT: Global issues and local concerns. In Ema Ushioda (ed.), *International perspectives on motivation*: *Language learning and professional challenges*. Houndsmill, Basingstoke: Palgrave Macmillan, 1–17.

Van Buuren, Stef. 2012. *Flexible imputation of missing data*. Boca Raton, FL: Chapman & Hall/CRC.

Van Buuren, Stef & Groothuis-Oudshoorn, Karin. 2011. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software* 45(3), 1–67.

Van Zeeland, Hilde. 2014. Lexical inferencing in first and second language listening. *The Modern Language Journal* 98(4), 1006–1021.

Van Zeeland, Hilde. 2017. Christopher Brumfit thesis award winner 2014 – Hilde van Zeeland: Four studies on vocabulary knowledge in and from listening: Findings and implications for future research. *Language Teaching* 50(1), 143–150.

Van Zeeland, Hilde & Schmitt, Norbert. 2013a. Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System* 41(3), 609–624.

Van Zeeland, Hilde & Schmitt, Norbert. 2013b. Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics* 34(4), 457–479.

Vasilyev, Edith. 2012. *Wie spricht Österreich? Übersicht über Migrantensprachen in Österreich* [How does Austria speak? Overview of migrant languages in Austria]. Wien: Österreichischer Integrationsfonds [Austrian Integration Fund]. http://www.forschungsnetzwerk.at/downloadpub/2012_n24_Dossier_Migrantensprachen_in_austria.pdf (9 May, 2018).

VERBI Software. 2017. *MAXQDA 2018* (Version 18.0.2). Berlin: VERBI Software. www.maxqda.com.

Verein Arbeitsgemeinschaft Media-Analysen [Association Working Group Media Analysis]. 2017. *Media-Analyse 2017*: *Internet* [Media analysis 2017: Internet]. Wien. https://www.media-analyse.at/table/3007 (6 June, 2018).

Verspoor, Marjolijn H.; De Bot, Kees & Van Rein, Eva. 2011. English as a foreign language: The role of out-of-school language input. In Annick De Houwer & Antje Wilton (eds.), *English in Europe today*: *Sociocultural and educational perspectives*. Amsterdam: John Benjamins, 147–166.

Vidal, Karina. 2003. Academic listening: A source of vocabulary acquisition? *Applied Linguistics* 24(1), 56–89.

Vidal, Karina. 2011. A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning* 61(1), 219–258.

Vilkaitė, Laura. 2017. Incidental acquisition of collocations in L2: Effects of adjacency and prior vocabulary knowledge. *ITL - International Journal of Applied Linguistics* 168(2), 248–277.

References

Voss, Erik. 2012. *A validity argument for score meaning of a computer-based ESL academic collocational ability test based on a corpus-driven approach to test design.* Ames, IA: Iowa State University, PhD thesis.

Wagner, Johannes. 2015. Designing for language learning in the wild: Creating social infrastructures for second language learning. In Teresa Cadierno & Søren W. Eskildsen (eds.), *Usage-based perspectives on second language learning.* Berlin: De Gruyter Mouton, 75–101.

Wagner, Wolfgang; Helmke, Andreas & Rösner, Ernst. 2009. *Deutsch Englisch Schülerleistungen International*: *Dokumentation der Erhebungsinstrumente für Schülerinnen und Schüler, Eltern und Lehrkräfte* [German English student achievements international: Documentation of research instruments for students, parents and teachers]. Frankfurt, Main: GFPF, DIPF.

Walker, David A. 2003. JMASM9: Converting Kendall's Tau for correlational or meta-analytic analyses. *Journal of Modern Applied Statistical Methods* 2(2), 525–530.

Waller, Gregor; Willemse, Isabel; Genner, Sarah; Suter, Lilian & Süss, Daniel. 2016. *JAMES Jugend, Aktivitäten, Medien - Erhebung Schweiz*: *Ergebnisbericht zur JAMES-Studie 2016* [JAMES youth, activities, media - Survey Switzerland: Results of the JAMES study 2016]. Zürich: Zürcher Hochschule für Angewandte Wissenschaften. https://www.zhaw.ch/de/psychologie/forschung/medienpsychologie/mediennutzung/james/#c77096 (6 June, 2018).

Walters, JoDee. 2012. Aspects of validity of a test of productive vocabulary: Lex30. *Language Assessment Quarterly* 9(2), 172–185.

Wang, Danping. 2012. Self-directed English language learning through watching English television drama in China. *Changing English* 19(3), 339–348.

Ward, Jeremy & Chuenjundaeng, Jitlada. 2009. Suffix knowledge: Acquisition and applications. *System* 37(3), 461–469.

Waring, Rob. 1997. A comparison of the receptive and productive vocabulary sizes of some second language learners. *Immaculata; The occasional papers at Notre Dame Seishin University* 1, 53–68.

Waring, Rob & Nation, I.S.P. 2004. Second language reading and incidental vocabulary learning. *Angles on the English-speaking World* 4, 11–23.

Waring, Rob & Takaki, Misako. 2003. At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language* 15(2), 130–163.

Webb, Stuart. 2005. Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition* 27(1), 33–52.

Webb, Stuart. 2007. The effects of repetition on vocabulary knowledge. *Applied Linguistics* 28(1), 46–65.

Webb, Stuart. 2008. The effects of context on incidental vocabulary learning. *Reading in a Foreign Language* 20(2), 232–245.

Webb, Stuart. 2010. A corpus driven study of the potential for vocabulary learning through watching movies. *International Journal of Corpus Linguistics* 15(4), 497–519.

Webb, Stuart. 2014. Repetition in incidental vocabulary learning. In Carol A. Chapelle (ed.), *The encyclopedia of applied linguistics*. Oxford, UK: Blackwell, n.d.

Webb, Stuart. 2015. Extensive viewing: Language learning through watching television. In David Nunan & Jack C. Richards (eds.), *Language learning beyond the classroom*. New York: Routledge, 159–168.

Webb, Stuart & Chang, Anna C.-S. 2012. Second language vocabulary growth. *RELC Journal* 43(1), 113–126.

References

Webb, Stuart & Chang, Anna C.-S. 2015. How does prior word knowledge affect vocabulary learning progress in an extensive reading program? *Studies in Second Language Acquisition* 37(4), 651–675.

Webb, Stuart & Macalister, John. 2013. Is text written for children useful for L2 extensive reading? *TESOL Quarterly* 47(2), 300–322.

Webb, Stuart; Newton, Jonathan & Chang, Anna. 2013. Incidental learning of collocation. *Language Learning* 63(1), 91–120.

Webb, Stuart & Piasecki, Anna. 2018. Re-examining the effects of word writing on vocabulary learning. *ITL - International Journal of Applied Linguistics* 169(1), 72–94.

Webb, Stuart & Rodgers, Michael P. H. 2009a. The lexical coverage of movies. *Applied Linguistics* 30(3), 407–427.

Webb, Stuart & Rodgers, Michael P. H. 2009b. Vocabulary demands of television programs. *Language Learning* 59(2), 335–366.

Webb, Stuart; Sasao, Yosuke & Ballance, Oliver. 2017. The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL - International Journal of Applied Linguistics* 168(1), 33–69.

Weber, Martina. 2008. Fremdsprachen in österreichischen Großunternehmen - eine Bedarfsanalyse [Foreign languages in large Austrian enterprises - a needs analysis]. In Sabine Tritscher-Archan (ed.), *Fremdsprachen für die Wirtschaft*: *Analysen, Zahlen, Fakten.* [Foreign languages for the economy: Analyses, figures, facts]. Vienna: Institut für Bildungsforschung der Wirtschaft, 147–170.

Weir, Cyril. 2005. *Language testing and validation*: *An evidence-based approach*. Basingstoke: Palgrave Macmillan.

Weltens, Bert & Grendel, Marjon. 1993. Attrition of vocabulary knowledge. In Robert Schreuder & Bert Weltens (eds.), *The bilingual lexicon*. Amsterdam: John Benjamins Publishing Company, 135–156.

Werquin, Patrick. 2016. International perspectives on the definition of informal learning. In Matthias Rohs (ed.), *Handbuch informelles Lernen.* [Handbook informal learning]. Wiesbaden: Springer, 39–64.

Wesche, Marjorie & Paribakht, Tahereh. 2010. *Lexical inferencing in a first and second language*: *Cross-linguistic dimensions*. Bristol: Multilingual Matters.

Westfall, Jacob; Kenny, David A. & Judd, Charles M. 2014. Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General* 143(5), 2020–2045.

Widdowson, Henry G. 1994. The ownership of English. *TESOL Quarterly* 28(2), 377–389.

Wieland, Alexandra. 2016. *Extramural English - Eine Untersuchung zur Wirkung von außerschulischem Fremdsprachenerwerb* [Extramural English - A study on the effects of out-of-school language acquisition]. Salzburg: PH Salzburg, BA thesis.

Wode, Henning. 1999. Incidental vocabulary acquisition in the foreign language classroom. *Studies in Second Language Acquisition* 21, 243–258.

World Health Organization. 2018. *Adolescent health.* http://www.who.int/topics/adolescent_health/en/ (6 June, 2018).

Wray, Alison. 2002. *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Wray, Alison. 2012. What do we (think we) know about formulaic language? An evaluation of the current state of play. *Annual Review of Applied Linguistics* 32, 231–254.

Wray, Alison. 2013. Formulaic language. *Language Teaching* 46(3), 316–334.

Wyn, Johanna & Cahill, Helen (eds.). 2015. *Handbook of children and youth studies*. Singapore: Springer.

References

Wysocki, Katherine & Jenkins, Joseph R. 1987. Deriving word meanings through morphological generalization. *Reading Research Quarterly* 22(1), 66–81.

Yap, Set-lee S. 1998. *Out-of-class use of English by secondary school students in a Hong Kong Anglo-Chinese school.* Hong Kong: University of Hong Kong, MA thesis.

Yeşilbursa, Amanda & Bilican, Recep. 2013. Validation of self-regulatory capacity in vocabulary learning scale in Turkish. *Procedia - Social and Behavioral Sciences* 70, 882–886.

Yi, Youngjoo. 2005. Asian adolescents' out-of-school encounters with English and Korean literacy. *Journal of Asian Pacific Communication* 15(1), 57–77.

Zechmeister, Eugene B.; Chronis, Andrea M.; Cull, William L.; D'Anna, Catherine A. & Healy, Noreen A. 1995. Growth of a functionally important lexicon. *Journal of Reading Behavior* 27(2), 201–212.

Zhang, Yining; Lin, Chin-Hsi; Zhang, Dongbo & Choi, Yunjeong. 2017. Motivation, strategy, and English as a foreign language vocabulary learning: A structural equation modelling study. *The British journal of educational psychology* 87(1), 57–74.

Zheng, Dongping; Bischoff, Michael & Gilliland, Betsy. 2015. Vocabulary learning in massively multiplayer online games: Context and action before words. *Educational Technology Research and Development* 63(5), 771–790.

Zheng, Yongyan. 2012. Exploring long-term productive vocabulary development in an EFL context: The role of motivation. *System* 40(1), 104–119.

Zichtl, Carolina. 2017. *A study of receptive and productive lexical knowledge of Austrian EFL students in year 8, 10, and 12.* Vienna: University of Vienna, MA thesis.

Zou, Di. 2017. Vocabulary acquisition through cloze exercises, sentence-writing and composition-writing: Extending the evaluation component of the Involvement Load Hypothesis. *Language Teaching Research* 21(1), 54–75.

Zuzanek, Jiri. 2005. Adolescent time use and well-being from a comparative perspective. *Loisir et Société / Society and Leisure* 28(2), 379–423.

Appendix A

Software used for analysis

Informed consent forms for students and parents used in the quantitative strand

Informed consent form for parents used in the qualitative strand

Extramural English Questionnaire

Information about the EEOLD for students

Extramural English Online Language Diary

Lex30

V_YesNo and translation items

Table A.1: Table of responses on the translation task accepted as correct and rejected as incorrect

Interview guide for focus group interviews (in German)

Visual material for the focus group interviews

Transcription conventions

Codebook used in the qualitative content analysis

Software used for analysis

Main software programmes

AntWordProfiler (Version 1.4.0w)	Anthony, Laurence. 2013. *AntWordProfiler* (1.4.0). Tokyo: Waseda University. www.laurenceanthony.net/software/antwordprofiler/.
MAXQDA 2018 (Version 18.0.2.)	VERBI Software. 2017. *MAXQDA 2018* (Version 18.0.2). Berlin: VERBI Software. https://www.maxqda.com.
Qualtrics survey software	Qualtrics LLC. 2016. *Qualtrics survey software*. Provo, Utah.
R (Version 3.5.1)	R Development Core Team. 2018. *R*: *A language and environment for statistical computing* (3.5.1). Vienna: The R Foundation for Statstical Computing. www.R-project.org/.
RStudio (Version 1.2.1335)	RStudio Team. 2018. *RStudio*: *Integrated Development for R* (1.2.1335). Boston, MA: RStudio, Inc. www.rstudio.com/.
SPSS (Version 22.0)	IBM Corp. 2016. *IBM SPSS Statistics for Windows* (22.0). Armonk, NY: IBM Corp.

R packages used

boot	Canty, Angelo & Ripley, Brian. 2017. *boot: Bootstrap R (S-Plus) functions.* (R package version 1.3-20).
bootES	Kirby, Kris N. & Gerlanc, Daniel. 2013. BootES: An R package for bootstrap confidence intervals on effect sizes. *Behavior research methods* 45(4), 905–927.
car	Fox, John & Weisberg, Sanford. 2011. *car: An {R} Companion to applied regression (2nd edition)*. Thousand Oaks, CA: SAGE.
compute.es	Del Re, A. C. 2013. *compute.es: Compute effect sizes* (R package version 0.2-2).
dplyr	Wickham, Hadley; François, Romain; Henry, Lionel & Müller, Kirill. 2018. *dplyr: A grammar of data manipulation* (R package version 0.7.8).
GGally	Schloerke, Barret; Crowley, Jason; Di Cook; Briatte, Francois; Marbach, Moritz; Thoen, Edwin; Elberg, Amos & Larmarange, Joseph. 2018. *GGally: Extension to 'ggplot2'* (R package version 1.4.0).
ggfortify	Tang, Yuan; Horikoshi, Masaaki & Li, Wenxuan. 2016. *ggfortify: Unified interface to visualize statistical result of popular R packages*: The R Journal 8(2), 478-489.

ggplot2	Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis*. New York: Springer.
Hmisc	Harrell, Frank E. Jr with contributions from Charles Dupont and many others. 2018. *Hmisc: Harrell Miscellaneous* (R package version 4.1-1).
lattice	Sarkar, Deepayan. 2008. *Lattice: Multivariate data visualization with R* . New York: Springer.
lme4	Bates, Douglas; Maechler, Martin; Bolker, Ben & Walker, Steve. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1), 1–48.
MuMIn	Barton, Kamil. 2018. *MuMIn: Multi-model inference* (R package version 1.42.1.).
nortest	Gross, Juergen & Ligges, Uwe. 2015. *nortest: Tests for normality* (R package version 1.0-4).
NSM3	Schneider, Grant; Chicken, Eric & Becvarik, Rachel. 2018. *NSM3: Functions and datasets to accompany Hollander, Wolfe, and Chicken - Nonparametric Statistical Methods, 3rd edition* (R package version 1.12.).
plyr	Wickham, Hadley. 2011. The split-apply-combine strategy for data analysis. *Journal of Statistical Software* 40(1), 1–29.
ppcor	Kim, Seongho. 2015. *ppcor: Partial and semi-partial (part) correlation* (R package version 1.1).
psych	Revelle, W. *psych: Procedures for personality and psychological research* (Version 1.8.12.). Evanston, Illinois: Northwestern University.
relaimpo	Grömping, Ulrike. 2006. Relative importance for linear regression in R: The package relaimpo. *Journal of Statistical Software* 17(1), 1–27.
reshape/ reshape2	Wickham, Hadley. 2007. Reshaping data with the reshape package. *Journal of Statistical Software* 21(12), 1–20.

Englisch in der Freizeit
Informationen für Schülerinnen und Schüler

Liebe Schülerin, lieber Schüler!

Im Rahmen meines Doktoratsstudiums an der Universität Wien führe ich ein Forschungsprojekt durch, bei dem es darum geht, herauszufinden, ob und wie Jugendliche in Österreich in ihrer Freizeit mit Englisch in Kontakt kommen, und ob die Beschäftigung mit Englisch einen Einfluss auf den Wortschatz hat. Deshalb brauche ich deine Unterstützung!

Deine Klasse wurde zur Teilnahme an diesem Projekt ausgewählt. Das bedeutet, dass ich euch, wenn ihr einverstanden seid, in diesem Schuljahr zweimal im Unterricht besuchen werde, um verschiedene Teile der wissenschaftlichen Studie durchzuführen.

In den verschiedenen Teilen des Forschungsprojekts

- wirst du in einem Fragebogen zu deinen Aktivitäten befragt
- werden mit Hilfe eines kurzen Online Sprachtagebuchs deine Gewohnheiten erhoben
- wird die Größe deines Wortschatzes mit Hilfe von verschiedenen Messinstrumenten geschätzt
- werden einige von euch eventuell in einem Interview genauer zu ihren Meinungen befragt

Die Teilnahme an diesem Projekt ist <u>freiwillig</u>. Die Daten in dieser Untersuchung sind <u>anonym</u> und werden ausschließlich für Forschungszwecke verwendet. Außerdem werden sie sicher aufbewahrt, sodass niemand außer mir Zugang dazu hat. Dies bedeutet auch, dass deine LehrerInnen die Ergebnisse nicht sehen werden und diese <u>keinen Einfluss auf deine Schulnoten</u> haben können.

Durch deine Mitarbeit kannst du einen wichtigen Beitrag zur Forschung leisten und du kannst dazu beitragen, dass WissenschaftlerInnen und EnglischlehrerInnen das Leben von Jugendlichen besser verstehen. Daher freue ich mich, wenn du an diesem wichtigen Forschungsprojekt teilnimmst und bitte um deine Unterstützung!

Wenn du noch Fragen hast, kannst du sie mir gleich stellen, oder auch jederzeit per E-Mail. Du kannst auch nach Ende der Datenerhebung Einsicht in deine Daten nehmen, oder dein Einverständnis zurückziehen und die Daten löschen lassen. Dies ist <u>bis Ende des Schuljahres 2016/17</u> (30. Juni 2017) jederzeit möglich. Wenn du deine Daten sehen oder löschen lassen möchtest, schreibe bitte eine E-Mail an <u>marlene.schwarz@univie.ac.at</u> oder rufe mich unter ███ ████████ an.

Nach dem Ende des Projektes und Auswertung der Daten (voraussichtlich im Schuljahr 2017/18), werde ich euch gerne über die Ergebnisse informieren, wenn eure Klasse das möchte.

Herzlichen Dank im Voraus für deine Mithilfe!

Mag. Marlene Schwarz

Institut für Anglistik und Amerikanistik, Universität Wien

Kontakt: <u>marlene.schwarz@univie.ac.at</u> oder per Telefon: ███████████

Englisch in der Freizeit
Informationen für Erziehungsberechtigte

Sehr geehrte Damen und Herren,

Im Rahmen meines Doktoratsstudiums führe ich ein Forschungsprojekt durch, bei dem es darum geht, herauszufinden, ob und wie Jugendliche in Österreich in ihrer Freizeit mit Englisch in Kontakt kommen und ob die Beschäftigung mit Englisch einen Einfluss auf den englischen Wortschatz hat. Dieses Projekt wird von der Universität Wien finanziert und wurde von Stadtschulrat sowie von der Direktion der Schule Ihres Kindes genehmigt.

Die Klasse Ihres Sohnes/Ihrer Tochter wurde zur Teilnahme an diesem Projekt ausgewählt. Dies bedeutet, dass ich sie in diesem Schuljahr zweimal im Unterricht besuchen werde, um die verschiedenen Teile der wissenschaftlichen Studie durchzuführen.

In den verschiedenen Teilen des Forschungsprojekts

- werden die SchülerInnen in einem Fragebogen zu ihren Aktivitäten befragt
- werden mit Hilfe eines kurzen Sprachtagebuchs ihre Gewohnheiten erhoben
- wird die Größe des Wortschatzes mit Hilfe von verschiedenen Messinstrumenten geschätzt
- werden einige SchülerInnen in einem Interview eventuell genauer zu ihren Meinungen befragt (*Falls Ihr Sohn/Ihre Tochter an so einem Interview teilnehmen möchte, werden Sie noch einmal gesondert um Ihre Zustimmung gefragt*.)

Die Teilnahme an diesem Projekt ist freiwillig. Die Daten, die in dieser Untersuchung erhoben werden, sind anonym und werden ausschließlich für Forschungszwecke verwendet. Dies bedeutet selbstverständlich auch, dass die Daten sicher aufbewahrt werden, keine Daten an Dritte weitergegeben werden, und die Rohdaten nach Abschluss des Projektes vernichtet werden. Selbstverständlich können daher die Ergebnisse dieses Forschungsprojektes keinen Einfluss auf die Schulnoten Ihres Kindes haben.

Ihr Sohn/Ihre Tochter wurde von mir darüber informiert, dass er/sie durch seine/ihre Mitarbeit einen wichtigen Beitrag zur Forschung leistet und dazu beitragen kann, dass WissenschaftlerInnen und EnglischlehrerInnen das Leben von Jugendlichen besser verstehen.

Falls er/sie zu einem späteren Zeitpunkt Einsicht in die Daten nehmen möchte, oder sein/ihr Einverständnis zurückziehen und die Daten löschen lassen möchte, ist dies bis Ende des Schuljahres 2016/17 (30. Juni 2017) jederzeit möglich. In diesem Fall schreiben Sie oder Ihr Kind bitte eine E-Mail an marlene.schwarz@univie.ac.at oder nehmen telefonischen Kontakt auf (█ █████████).

Ich bitte Sie den beiliegenden Abschnitt in Absprache mit Ihrem Kind auszufüllen und zu unterschreiben, und dem/der KlassenlehrerIn zukommen zu lassen.

Falls Sie Fragen haben, stehe ich gerne jederzeit per E-Mail zur Verfügung.

Ich bedanke mich für Ihre Unterstützung und verbleibe mit freundlichen Grüßen,

Mag. Marlene Schwarz

Institut für Anglistik und Amerikanistik, Universität Wien

Kontakt: marlene.schwarz@univie.ac.at oder per Telefon: ██████████

Appendix A

✂ -

Mein Sohn / Meine Tochter _____, Schüler/in der ____ Klasse,

☐ nimmt mit meiner Erlaubnis am Forschungsprojekt „Englisch in der Freizeit" teil.

Ich erkläre mich damit einverstanden, dass die Daten meines Kindes zu den im Informationsblatt genannten Forschungszwecken erhoben, verarbeitet und genutzt werden dürfen.

Mein Kind und ich wurden darauf hingewiesen, dass die Erhebung und Verarbeitung dieser Daten auf freiwilliger Basis erfolgt und dass diese Einwilligung bis Ende des Schuljahres 2016/17 (30. Juni 2017) jederzeit ohne Angabe von Gründen zurückgezogen und die Löschung der Daten verlangt werden kann.

☐ nimmt nicht am Forschungsprojekt „Englisch in der Freizeit" teil.

_____

Datum   und   Unterschrift   des/der
Erziehungsberechtigten

Englisch in der Freizeit
Informationen für Erziehungsberechtigte
Gruppeninterview

Sehr geehrte Damen und Herren,

Wie Sie bereits im Laufe dieses Schuljahres informiert wurden, führe ich im Rahmen meines Doktoratsstudiums an der Universität Wien ein Forschungsprojekt durch, bei dem es darum geht, herauszufinden, ob und wie Jugendliche in Österreich in ihrer Freizeit mit Englisch in Kontakt kommen und ob die Beschäftigung mit Englisch einen Einfluss auf den englischen Wortschatz hat.

Ihr Sohn/Ihre Tochter hat am ersten Teil dieses Projektes teilgenommen. Er/Sie hat sich nun auch freiwillig dazu bereit erklärt, am zweiten Teil dieses Projektes teilzunehmen. Dabei soll ein Gruppeninterview (ca. 1 Stunde) mit 5 SchülerInnen aus der Schule Ihres Kindes durchgeführt werden.

In diesem Gruppeninterview werden die SchülerInnen zu folgenden Themen diskutieren
- Typische Beschäftigungen mit Englisch in der Freizeit
- Gründe für die Beschäftigung mit Englisch in der Freizeit
- Mögliche Lerneffekte
- Unterschiede zwischen Englisch in- und außerhalb der Schule

Die Teilnahme an diesem Gruppeninterview ist freiwillig. Damit die Daten ausgewertet werden können, müssen sie auf Tonband aufgenommen werden (nur Audio, kein Video). Selbstverständlich werden auch diese Daten anonymisiert und ausschließlich für Forschungszwecke verwendet. Dies bedeutet selbstverständlich auch, dass die Daten sicher aufbewahrt werden, nur zu Forschungszwecken verwendet werden, und die Rohdaten nach Abschluss des Projektes vernichtet werden.

Das Gruppeninterview wird möglicherweise außerhalb der Unterrichtszeit stattfinden, den genauen Termin kann ich erst koordinieren, wenn alle TeilnehmerInnen feststehen. Es wird selbstverständlich darauf geachtet, den Aufwand für die SchülerInnen so gering wie möglich zu halten.

Falls Ihr Sohn/Ihre Tochter zu einem späteren Zeitpunkt sein/ihr Einverständnis zurückziehen und die Daten löschen lassen möchte, ist dies bis Ende des Schuljahres 2016/17 (30. Juni 2017) jederzeit möglich. In diesem Fall schreiben Sie oder Ihr Kind bitte eine E-Mail an marlene.schwarz@univie.ac.at oder nehmen telefonischen Kontakt auf (███████████).

Damit Ihr Sohn/Ihre Tochter am Gruppeninterview teilnehmen darf, wird Ihre Einwilligung zu der Tonbandaufnahme benötigt. Ich bitte Sie den beiliegenden Abschnitt in Absprache mit Ihrem Kind auszufüllen und zu unterschreiben.

Falls Sie Fragen haben, stehe ich gerne jederzeit per E-Mail zur Verfügung.

Ich bedanke mich für Ihre Unterstützung und verbleibe mit freundlichen Grüßen,

Mag. Marlene Schwarz

Institut für Anglistik und Amerikanistik, Universität Wien

Kontakt: marlene.schwarz@univie.ac.at

Appendix A

✂ -

Mein Sohn / Meine Tochter _____, Schüler/in der ____ Klasse,

☐ darf mit meiner Erlaubnis am Gruppeninterview für das Forschungsprojekt „Englisch in der Freizeit" teilnehmen und auf Tonband aufgezeichnet werden.

Ich erkläre mich damit einverstanden, dass mein Kind zu den im Informationsblatt genannten Forschungszwecken auf Tonband aufgezeichnet werden darf und diese Daten in anonymisierter Form genutzt werden dürfen.

Mein Kind und ich wurden darauf hingewiesen, dass die Erhebung und Verarbeitung dieser Daten auf freiwilliger Basis erfolgt und dass diese Einwilligung bis Ende des Schuljahres 2016/17 (30. Juni 2017) jederzeit ohne Angabe von Gründen zurückgezogen und die Löschung der Daten verlangt werden kann.

☐ darf nicht am Gruppeninterview für das Forschungsprojekt „Englisch in der Freizeit" teilnehmen.

_____

Datum und Unterschrift des/der Erziehungsberechtigten

Englisch
in der Freizeit

Liebe Schülerin, lieber Schüler!

Dieser Fragebogen ist der erste Teil der Studie zu Englisch in der Freizeit. Die Beantwortung dauert ca. 40 Minuten.

Zur Erinnerung: Deine Antworten sind **anonym** und werden streng vertraulich behandelt. Es ist kein Test, also beantworte die Fragen bitte **alleine, spontan und ehrlich** – es gibt keinerlei falsche Antworten! Mich interessiert, was **du** wirklich machst und denkst.

Damit ich deinen Fragebogen den anderen Teilen der Studie zuordnen kann, bitte ich dich hier deinen persönlichen Code gut leserlich in BLOCKBUCHSTABEN einzufüllen:

Mein Code:	☐☐ 2. und 3. Buchstabe deines Vornamens
	☐ der letzte Buchstabe des Vornamens deiner Mutter
	☐☐ Geburtsmonat (z.B. 03 für März)

Hinweise zum Ausfüllen des Fragebogens:

Bei den meisten Fragen musst du dich nur für eine Antwortmöglichkeit entscheiden und …

….ein Kästchen ankreuzen

Hast du Geschwister?   ☒ Ja   ☐ Nein

…die zutreffende Zahl ankreuzen

Wie oft isst du Erdbeeren?	Fast nie	Selten	Oft	Sehr oft
	1	✗	3	4

Bei allen Fragen, die mit ➡ *gekennzeichnet sind, kannst du eine Antwort in eigenen Worten geben.*

Bitte schreib leserlich!

Wie findest du Clowns? ➡ *Ich finde Clowns total witzig.*

➢ Falls du etwas falsch angekreuzt oder hingeschrieben hast, streich es bitte klar durch!

➢ Bei manchen Fragen stehen genauere Hinweise in *kursiv* angeben. Bitte **lies genau**, was du machen sollst.

Falls du Fragen hast, wende dich bitte einfach an mich!
Danke für deine Bereitschaft, nun legen wir los!

1a. Wie viele der folgenden Dinge habt ihr zu Hause?	**0**	**1**	**2**	**3 oder mehr**
Handys	☐	☐	☐	☐
Fernseher	☐	☐	☐	☐
Computer (Laptops, Netbooks und Standcomputer)	☐	☐	☐	☐

1b. Welche von diesen Geräten kannst du zu Hause benutzen?
➢ *Bitte kreuze **alle** Geräte an, die du zu Hause benutzen kannst!*

☐ Tablet/iPad	☐ Spielekonsole (z.B. Wii, Playstation)
☐ Radio/Hi-Fi Anlage	☐ DVD-Player (oder Blu-ray Player)
☐ Mp3-Player (iPod oder andere)	☐ E-Book-Reader (z.B. Kindle)

1c. Hast du …	**Ja**	**Nein**
ein eigenes Smartphone?	☐	☐
einen eigenen Computer oder Laptop?	☐	☐
einen eigenen Mp3-Player (iPod oder andere)?	☐	☐
ein eigenes Zimmer?	☐	☐

1d. Gibt es bei dir zu Hause einen Internetzugang?   ☐ Ja   ☐ Nein

1

1e. Auf welchen Geräten benutzt du normalerweise das Internet?

1. ➡ _____ 2. ➡ _____ 3. ➡ _____

1f. Benutzt du das Internet täglich?   ☐ Ja   ☐ Nein

⇨ wenn **ja**, wie viele Stunden pro Tag? Ca. _____ Stunden

1g. Welche fünf Webseiten besuchst du am häufigsten?

1. ➡ _____ 2. ➡ _____ 3. ➡ _____

4. ➡ _____ 5. ➡ _____

1h. Wie oft machst du diese Tätigkeiten in deiner Freizeit allgemein?	(Fast) nie	Ein paar Mal pro Jahr	Ein paar Mal pro Monat	Ein paar Mal pro Woche	(Fast) täglich
Musik hören	1	2	3	4	5
Radio hören	1	2	3	4	5
Fernsehen	1	2	3	4	5
Filme oder Serien auf DVD ansehen	1	2	3	4	5
Filme oder Serien im Internet ansehen	1	2	3	4	5
soziale Netzwerke benutzen (z.B. Facebook, Instagram,…)	1	2	3	4	5
Videos im Internet sehen	1	2	3	4	5
Spiele alleine spielen (am Computer, Spielekonsolen oder online)	1	2	3	4	5
Spiele mit anderen gemeinsam spielen (z.B. Multiplayer Online Games)	1	2	3	4	5
Spiele am Handy oder Tablet spielen	1	2	3	4	5
Zeitungen oder Zeitschriften lesen (auch online)	1	2	3	4	5
Bücher oder E-Books lesen	1	2	3	4	5
Hörbücher hören	1	2	3	4	5
FreundInnen treffen	1	2	3	4	5
Sport machen	1	2	3	4	5
Musik machen (singen oder ein Instrument spielen)	1	2	3	4	5
auf Konzerte gehen	1	2	3	4	5
ins Kino gehen	1	2	3	4	5
ins Theater gehen	1	2	3	4	5
andere Aktivitäten, nämlich: ➡ _____	1	2	3	4	5

1i. Denk bitte an deinen ganz normalen Alltag in Österreich:
   Wo begegnet dir Englisch am meisten?
   ➢*Bitte zähle deine Top 3 auf und beginne mit dem wichtigsten!*

1. ➡ _____

2. ➡ _____

3. ➡ _____

2. Englische Freizeitaktivitäten

Mich interessiert, was du **in deiner Freizeit** (=außerhalb des Schulunterrichts) **auf Englisch** machst.
➢ *Es geht hier nur um Kontakte mit Englisch in deiner Freizeit, die du <u>nicht</u> für den Schulunterricht machen musst (also <u>keine</u> Hausübungen, Referatsvorbereitungen, Leseaufgaben …)!*

2a. Wie oft machst du diese Tätigkeiten <u>in deiner Freizeit</u> auf Englisch? ➢ *Fülle bitte auch die offenen Fragen aus!*	(Fast) nie	Ein paar Mal pro Jahr	Ein paar Mal pro Monat	Ein paar Mal pro Woche	(Fast) täglich
englischsprachige Musik					
…am Handy/mp3-Player hören	1	2	3	4	5
…auf CD oder im Radio hören	1	2	3	4	5
…auf Spotify oder anderen Musikdiensten hören	1	2	3	4	5
…Musikvideos im Internet ansehen	1	2	3	4	5
…mitsingen	1	2	3	4	5
…selbst singen (auch Karaoke)	1	2	3	4	5
…auf Konzerten hören	1	2	3	4	5
englischsprachige Songtexte					
…lesen	1	2	3	4	5
…übersetzen oder Übersetzungen lesen	1	2	3	4	5
…schreiben	1	2	3	4	5
englischsprachige Filme					
…im Fernsehen ansehen	1	2	3	4	5
…auf DVD / Blu-ray ansehen **mit** Untertiteln	1	2	3	4	5
…auf DVD / Blu-ray ansehen **ohne** Untertitel	1	2	3	4	5
…im Internet ansehen (z.B. Youtube, Netflix, …) **mit** Untertiteln	1	2	3	4	5
…im Internet ansehen (z.B. Youtube, Netflix, …) **ohne** Untertitel	1	2	3	4	5
…im Kino ansehen	1	2	3	4	5
englischsprachige Serien					
…im Fernsehen ansehen	1	2	3	4	5
…auf DVD / Blu-ray ansehen **mit** Untertiteln	1	2	3	4	5
…auf DVD / Blu-ray ansehen **ohne** Untertitel	1	2	3	4	5
…im Internet ansehen (z.B. Youtube, Netflix, …) **mit** Untertiteln	1	2	3	4	5
…im Internet ansehen (z.B. Youtube, Netflix, …) **ohne** Untertitel	1	2	3	4	5

⇨ In welcher Sprache sind die Untertitel bei Filmen und Serien meistens?

➡ _____

	(Fast) nie	Ein paar Mal pro Jahr	Ein paar Mal pro Monat	Ein paar Mal pro Woche	(Fast) täglich
englischsprachige Sendungen (z.B. Dokus, Reportagen etc.)					
…im Fernsehen ansehen	1	2	3	4	5
…im Radio anhören	1	2	3	4	5
…im Internet ansehen	1	2	3	4	5
englischsprachige Videoclips (z.B. Youtube,…)					
…im Internet ansehen	1	2	3	4	5
…selbst machen	1	2	3	4	5
englischsprachige Spiele					
…am Tablet, Handy, iPod spielen	1	2	3	4	5
…am Computer oder auf Spielekonsolen spielen	1	2	3	4	5
…mit anderen über Internet spielen (Multiplayer Online Games)	1	2	3	4	5
⇨ **Wenn ich mit anderen im Internet spiele,** ➢ *Lass diese Frage aus, wenn du nicht im Internet spielst!*					
…chatte ich auf Englisch in In-Game Chats	1	2	3	4	5
…spreche ich Englisch über TeamSpeak, Mumble oder ähnliche Dienste	1	2	3	4	5
englischsprachige Bücher					
…in Papierform lesen	1	2	3	4	5
…digital (z.B. E-Reader wie Kindle) lesen	1	2	3	4	5
…als Hörbuch anhören	1	2	3	4	5

Zur Erinnerung: **Denk bitte nur an deine *Freizeit!*** *(keine Hausübungen, Schulausflüge, …)*	(Fast) nie	Ein paar Mal pro Jahr	Ein paar Mal pro Monat	Ein paar Mal pro Woche	(Fast) täglich
englischsprachige Artikel (z.B. Zeitung, Zeitschriften, Nachrichtenportale etc.)					
…in Papierform lesen	1	2	3	4	5
…im Internet lesen	1	2	3	4	5
englischsprachige Informationstexte (z.B. Anleitungen, Rezepte etc.)					
…in Papierform lesen	1	2	3	4	5
…im Internet lesen (auch z.B. Wikipedia etc.)	1	2	3	4	5
englischsprachige Geschichten / Fan-Fictions					
…lesen	1	2	3	4	5
…schreiben	1	2	3	4	5
englischsprachige Comics / Mangas					
…lesen	1	2	3	4	5
…schreiben	1	2	3	4	5
englischsprachige Blogs / Foreneinträge (z.B. zu Spielen)					
…lesen	1	2	3	4	5
…schreiben	1	2	3	4	5
Suchmaschinen (z.B. Google) auf Englisch benutzen	1	2	3	4	5
englischsprachige Apps am Handy benutzen	1	2	3	4	5

⇨ welche Apps?

➡ _____

	(Fast) nie	Ein paar Mal pro Jahr	Ein paar Mal pro Monat	Ein paar Mal pro Woche	(Fast) täglich
englischsprachige E-Mails					
…lesen	1	2	3	4	5
…schreiben	1	2	3	4	5
in sozialen Netzwerken (z.B. Facebook, Twitter, Instagram)					
…englischsprachige Nachrichten lesen	1	2	3	4	5
…englischsprachige Nachrichten schreiben	1	2	3	4	5
…Statusmeldungen / Kommentare auf Englisch lesen	1	2	3	4	5
…Statusmeldungen / Kommentare (auch zu Fotos) auf Englisch schreiben	1	2	3	4	5
englischsprachige SMS / WhatsApp Nachrichten					
…lesen	1	2	3	4	5
…schreiben	1	2	3	4	5
auf Englisch chatten	1	2	3	4	5

⇨ mit wem schreibst du englische Nachrichten oder chattest du? *(Bitte keine Namen!)*

➡ _____

	(Fast) nie	Ein paar Mal pro Jahr	Ein paar Mal pro Monat	Ein paar Mal pro Woche	(Fast) täglich
Listen oder Notizen auf Englisch schreiben	1	2	3	4	5
Tagebuch auf Englisch schreiben	1	2	3	4	5
Englisch sprechen					
…via Skype oder ähnliche Internetdienste	1	2	3	4	5
…am Telefon	1	2	3	4	5
…persönlich	1	2	3	4	5

⇨ Mit wem sprichst du Englisch?

➡ _____

	(Fast) nie	Ein paar Mal pro Jahr	Ein paar Mal pro Monat	Ein paar Mal pro Woche	(Fast) täglich
auf Englisch denken / mit dir selbst reden	1	2	3	4	5
englische Wörter in anderen Sprachen verwenden	1	2	3	4	5
englische Theaterstücke					
…ansehen	1	2	3	4	5
…selbst spielen	1	2	3	4	5
Ich mache noch andere Aktivitäten auf Englisch, nämlich: ➡ _____	1	2	3	4	5

2b. *Bitte gib hier TITEL oder genauere BESCHREIBUNGEN an!*

a)	Hast du einen englischsprachigen Lieblingsfilm (oder mehrere)?	☐ Ja	☐ Nein

⇨ Wenn **ja**, welchen? ➡ _____

b)	Hast du eine englischsprachige Lieblingsserie (oder mehrere)?	☐ Ja	☐ Nein

⇨ Wenn **ja**, welche? ➡ _____

c)	Hast du ein englischsprachiges Lieblingscomputerspiel (oder mehrere)?	☐ Ja	☐ Nein

⇨ Wenn **ja**, welches? ➡ _____

d)	Hast du ein englischsprachiges Lieblingsbuch (oder mehrere)?	☐ Ja	☐ Nein

⇨ Wenn **ja**, welches? ➡ _____

e)	Liest du etwas Bestimmtes auf Englisch (z.B. Zeitschrift, Blog,…)? *Bitte beschreibe es!*	☐ Ja	☐ Nein

⇨ Wenn **ja**, was? ➡ _____

f)	Siehst du dir etwas Bestimmtes auf English im Internet an (z.B. Vlog,…)? *Bitte beschreibe es!*	☐ Ja	☐ Nein

⇨ Wenn **ja**, was? ➡ _____

g) Bei englischsprachiger Musik sind die Songtexte für mich…

☐ Gar nicht wichtig   ☐ Nicht so wichtig   ☐ Eher wichtig   ☐ Sehr wichtig

2c. Welche Sprachen verwendest du in deiner FREIZEIT (Freunde, Familie, Aktivitäten,…)**?**

1. ➡ _____   2. ➡ _____

3. ➡ _____   4. ➡ _____

2d. Denk bitte an <u>alle</u> Sprachen, die du in deiner FREIZEIT (also außerhalb des Unterrichts) verwendest
Wie viel Zeit verbringst du mit deinen Sprachen im Vergleich? (**Ohne HÜs, Lernen und Nachhilfe!**)
➤ *Teile den Kreis in Teile und schreibe dazu, für welche Sprache die Teile stehen!*

Dein Kreis:
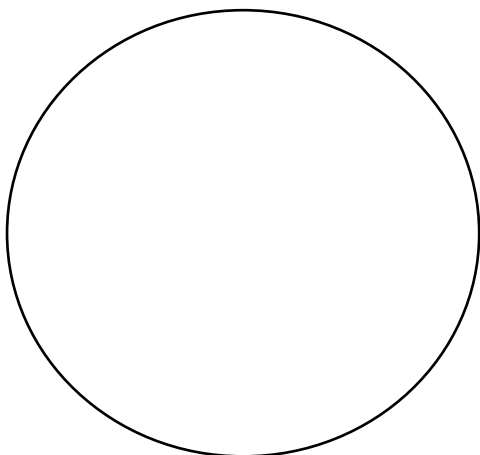
Beispiel 1:
Für einen Jugendlichen in Finnland könnte der Kreis zum Beispiel so aussehen:

Schwedisch

Englisch

Finnisch
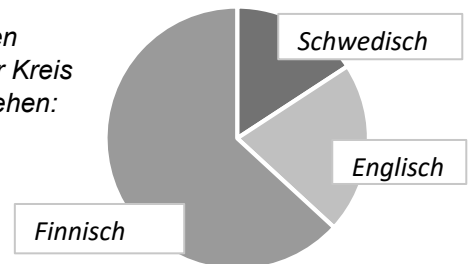
Beispiel 2:
Für ein 4-jähriges Kind in Griechenland könnte der Kreis zum Beispiel so aussehen:

Griechisch

5

2e. Warum verwendest du Englisch in deiner Freizeit?	Trifft gar nicht zu	Trifft eher nicht zu	Trifft eher zu	Trifft völlig zu
Vieles gibt es (zumindest vorübergehend) nur auf Englisch.	1	2	3	4
Ich brauche Englisch für internationale Kontakte und Freundschaften.	1	2	3	4
Ich möchte meine Englischkenntnisse verbessern.	1	2	3	4
Es interessiert mich, wie Dinge im Original sind (z.B. Filme, Bücher).	1	2	3	4
Vieles klingt auf Englisch besser.	1	2	3	4
Ich habe einfach Lust, Englisch zu verwenden	1	2	3	4

Ich habe andere Gründe, nämlich:

➡ _____

3. Meinungen zu Englisch

Hier sind einige Meinungen von Leuten zur Sprache Englisch. Wie sehr stimmst du diesen Ideen zu?

|	Stimme gar nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme völlig zu
Englisch ist für das Leben nach der Schule wichtig.	1	2	3	4
Wenn man viel reist, muss man Englisch können.	1	2	3	4
Englisch bereichert die deutsche Sprache.	1	2	3	4
Man kann auch ohne Englischkenntnisse international sein.	1	2	3	4
Englisch ist auch innerhalb von Österreich wichtig.	1	2	3	4
Im täglichen Leben von Jugendlichen hat Englisch in der Freizeit eine größere Bedeutung als in dem von Erwachsenen.	1	2	3	4
Englisch zu können ist für das Leben in Zukunft sehr wichtig.	1	2	3	4
Englischkenntnisse werden in Österreich überbewertet.	1	2	3	4
Englisch ist moderner als Deutsch.	1	2	3	4
Für die Generation 50+ ist Englisch nicht so wichtig wie für junge Leute.	1	2	3	4
Englisch zu beherrschen ist für eine gute Zukunft nicht nötig.	1	2	3	4
Es ist wichtig für Jugendliche, Englisch zu lernen.	1	2	3	4
Mit Englisch kann man sich überall verständlich machen.	1	2	3	4
Englisch ist für das Leben außerhalb der Schule wichtig.	1	2	3	4
Englisch stellt eine Bedrohung für die deutsche Sprache dar.	1	2	3	4
Gute Englischkenntnisse sind heutzutage in Österreich enorm wichtig.	1	2	3	4
Man kann sich auch ohne Englisch gut mit Leuten aus anderen Ländern verständigen.	1	2	3	4
Es ist wichtig Englisch zu können, um später einen guten Job zu bekommen.	1	2	3	4
Jugendliche können auch ohne Englisch gut zurechtkommen.	1	2	3	4
Englisch klingt schöner als Deutsch.	1	2	3	4
Für das Leben in Österreich braucht man kein Englisch.	1	2	3	4

4. Sprachliche Umgebung

4a. Wo siehst du oder hörst du Englisch im Alltag?

	(Fast) nie	Selten	Oft	Sehr oft	Da bin ich nie
auf der Straße	1	2	3	4	0
an öffentlichen Plätzen (z.B. Bahnhof, Parks, etc.)	1	2	3	4	0
in öffentlichen Verkehrsmitteln	1	2	3	4	0
in öffentlichen Ämtern, Banken etc.	1	2	3	4	0
in Geschäften oder Einkaufszentren	1	2	3	4	0
in Cafés, Restaurants etc.	1	2	3	4	0
an Orten, wo ich meine Hobbys ausübe (Fitnessstudio, Musikschule,..)	1	2	3	4	0
zu Hause	1	2	3	4	0

4b. Bitte denke jetzt an alle Situationen in denen du Englisch verwendest, also es sprichst, liest, hörst, oder schreibst. Denk auch an alle „unwichtigen" Situationen, z.B. wenn du nur ein paar Wörter verwendest.

Wo verwendest du Englisch mehr?　　☐ Im Schulunterricht　　☐ in meiner Freizeit

5. Strategien für neue englische Wörter

5a. Wie oft achtest du auf neue englische Wörter, wenn sie dir in der Freizeit begegnen?

	Trifft gar nicht zu	Trifft eher nicht zu	Trifft eher zu	Trifft völlig zu
Ich achte auf neue englische Wörter, wenn sie für den Inhalt wichtig sind.	1	2	3	4
Ich achte in englischen Medien bewusst auf Wörter, die ich nicht kenne.	1	2	3	4
Mich interessieren neue englische Wörter nicht – Hauptsache, ich verstehe um was es geht.	1	2	3	4
Mir fallen neue englische Wörter auf, obwohl ich nicht bewusst darauf achte.	1	2	3	4
Neue englische Wörter interessieren mich in meiner Freizeit nicht.	1	2	3	4

5b. Was machst du, wenn dir in deiner Freizeit ein englisches Wort begegnet, das du nicht kennst?

	(Fast) nie	Selten	Oft	Sehr oft
Ich überlege mir, welche Art von Wort das ist (Verb, Hauptwort…).	1	2	3	4
Ich suche in einem Wort nach Wortteilen, die ich kenne.	1	2	3	4
Ich überlege, ob es in anderen Sprachen, die ich kann, ein ähnliches Wort gibt.	1	2	3	4
Wenn es in einem Film oder einer Serie vorkommt, versuche ich die Bedeutung mit Hilfe des Bildes (und der Handlung) zu erraten.	1	2	3	4
Wenn es in einem Text vorkommt, versuche ich die Bedeutung aus dem Zusammenhang zu erraten.	1	2	3	4
Ich schaue in einem Wörterbuch nach (auch online oder am Handy).	1	2	3	4
Ich frage jemanden (Eltern, Geschwister, Freunde,…) was das Wort bedeutet.	1	2	3	4
Ich mache gar nichts.	1	2	3	4
Ich mache etwas anderes: ➡	1	2	3	4

5c. Wenn du eine Vermutung hast, was ein neues englisches Wort bedeutet, machst du dann noch etwas, um diese Vermutung zu überprüfen?　　☐ Ja　　☐ Nein

⇨ *Wenn* **ja**, *beschreibe bitte kurz was du machst*!

➡

5d. Machst du etwas Bestimmtes, um dir neue englische Wörter zu merken? ☐ Ja ☐ Nein
⇨ *Wenn **ja**, beschreibe bitte kurz was du machst!*

➡ _____

_____

_____

6. Sprachlicher Hintergrund

➢ *In einigen Fragen kommen die Begriffe „**englischsprachiges Land**" und „**nicht-englischsprachiges Land**"*
vor.
*Mit „**englischsprachigen Ländern**" sind Staaten gemeint, in denen Englisch eine offizielle Sprache ist (z.B.*
Großbritannien, USA, Australien, …).
*Mit „**nicht-englischsprachigen Ländern**" sind Staaten gemeint, in denen Englisch keine offizielle Sprache ist,*
das bedeutet aber natürlich nicht, dass man dort nicht Englisch sprechen kann.

6a. Welche Sprache sprichst du …
➢ *Wenn eine Frage nicht auf dich zutrifft, lass die Zeile*
daneben einfach leer!   ⬇ *Bitte trag hier die Sprachen ein!*

…mit deiner Mutter?	
…mit deinem Vater?	
…mit deinen Geschwistern?	
… mit anderen Menschen, die bei dir zu Hause wohnen? (z.B. Großeltern, Partner/in von Mutter oder Vater)	
…meistens mit deinen besten Freunden?	
Welche Sprache sprechen deine Eltern miteinander?	

6b. Gibt es Menschen, mit denen du (außerhalb des Unterrichts) regelmäßig
Englisch sprichst?   ☐ Ja ☐ Nein

⇨ *Wenn **ja**, mit wem?* ➡ _____

6c. Wann hast du begonnen Englisch zu lernen?   Mit ca. _____ Jahren.

6d. Lernst du in der Schule neben Englisch noch (eine) andere Fremdsprache(n)?   ☐ Ja ☐ Nein

⇨ *Wenn **ja**, welche?* ➡ _____

6e. Warst du auf Sprachcamps oder internationalen Feriencamps, auf denen
Englisch gesprochen wurde?   ☐ Ja ☐ Nein

⇨ *Wenn **ja**, wie oft und wie lange?* ➡ _____

6f. Wie oft warst du schon in englischsprachigen Ländern auf Urlaub?
☐ noch nie ☐ 1- bis 3-mal ☐ 4- bis 6-mal ☐ mehr als 6-mal
⇨ Und zwar in:
➡ _____

6g. Wie viel Englisch hast du auf vergangenen Urlauben in nicht-englischsprachigen Ländern verwendet?

☐ sehr wenig ☐ ziemlich wenig ☐ ziemlich viel ☐ sehr viel

6h. Hast du jemals mehr als 3 Monate in einem Land gewohnt, wo du dich auf Englisch verständigt hast? (auch nicht-englischsprachige Länder!)
☐ Ja ☐ Nein

⇨ Wenn **ja**, wo und wie lange (in Monaten)?

1. In ➡ _____ für ca. _____ Monate

2. In ➡ _____ für ca. _____ Monate

3. In ➡ _____ für ca. _____ Monate

6i. Wie schätzt du deine Englischkenntnisse ein?

➢ *Kreuze für jede der vier Fertigkeiten die Beschreibung an, die am besten auf dich zutrifft. Bitte kreuze immer nur eine Möglichkeit an.*

LESEN auf Englisch	Ich kann kurze, eher einfache Texte verstehen und Informationen z.B. in Speisekarten, Fahrplänen oder Werbeanzeigen finden. Ich kann auch persönliche Nachrichten von einem Freund verstehen.	☐
	Ich kann Informationen in Briefen, Werbebroschüren oder Gebrauchsanweisungen verstehen, und ich verstehe auch die Hauptaussagen von Artikeln oder Sachtexten über Themen, die mich interessieren.	☐
	Ich kann Artikel und Sachtexte zu vielen verschiedenen Themen verstehen, auch wenn ich mich mit dem Thema nicht gut auskenne. Ich kann auch Kurzgeschichten und Romane verstehen.	☐

HÖREN auf Englisch	Ich kann kurze Gespräche zu persönlichen Informationen z.B. über Hobbies und Berufe gut verstehen. Ich kann auch den Wetterbericht oder Wegbeschreibungen verstehen.	☐
	Ich kann Gespräche über Themen, die mich interessieren, gut verstehen. Ich kann auch in Fernseh- und Radiosendungen oder Filmen die wichtigsten Punkte verstehen.	☐
	Ich kann längeren Vorträgen, Berichten oder Interviews gut folgen, auch wenn ich das Thema nicht gut kenne. Ich kann die Fernsehnachrichten, Radiosendungen oder Filme gut verstehen.	☐

SCHREIBEN auf Englisch	Ich kann kurze, eher einfache Texte schreiben, ich kann z.B. ein Formular ausfüllen, eine kurze persönliche Nachricht verfassen oder eine Einladung schreiben.	☐
	Ich kann zusammenhängende Texte schreiben über Themen, die mich interessieren. Ich kann längere Nachrichten an Freunde verfassen und ich kann Ereignisse und Gefühle z.B. in einer Geschichte beschreiben.	☐
	Ich kann übersichtliche, detaillierte Texte über viele verschiedene Themen schreiben. Ich kann z.B. meine Meinung zu einer Neuigkeit ausdrücken oder Argumente für und gegen etwas in einem Artikel präsentieren.	☐

SPRECHEN auf Englisch	Ich kann in Alltagssituationen gut kommunizieren, z.B. Einkaufen gehen, Essen bestellen oder einen Weg beschreiben. Ich kann auch einiges über mich selbst erzählen und einfache Gespräche führen.	☐
	Ich kann mich auf Reisen in anderen Ländern gut verständigen. Ich kann auch Gespräche über Themen führen, die mich interessieren, über persönliche Erfahrungen berichten und meine Meinung äußern und begründen.	☐
	Ich kann mich zu vielen verschiedenen Themen ziemlich flüssig unterhalten und in Diskussionen meine Meinung klar darlegen und vertreten. Ich kann auch Präsentationen geben und z.B. die Vor- und Nachteile eines Themas erläutern.	☐

6j. Welche Jahresnote hattest du letztes Schuljahr im Fach Englisch?
☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

6k. Mit welcher Note denkst du, dass du dieses Schuljahr im Fach Englisch abschließen wirst?
☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

7. Du und deine Familie

Du hast es fast geschafft! Bitte beantworte zum Schluss nur noch einige Fragen, damit ich die Ergebnisse besser auswerten kann.
➤ *Wenn eine Frage nicht auf dich oder deine Eltern zutrifft, lass die Zeile einfach leer!*

7a. dein Geschlecht: ☐ weiblich   ☐ männlich

7b. dein Geburtsjahr: _ _ _ _   **und dein Geburtsland:** _____

7c. Was ist die <u>höchste</u> Ausbildung, die deine Eltern <u>abgeschlossen</u> haben?	Mutter	Vater
Er / Sie hat keine Ausbildung abgeschlossen.	☐	☐
Pflichtschule (Volksschule + Hauptschule oder Unterstufe)	☐	☐
Berufsschule / Lehre	☐	☐
Lehre mit Meisterprüfung	☐	☐
Höhere Schule ohne Matura (z.B. Fachschule, Handelsschule, Krankenpflegeschule,..)	☐	☐
Allgemeinbildende Höhere Schule mit Matura (AHS)	☐	☐
Berufsbildende Höhere Schule mit Matura (BHS, z.B. HAK, HTL, HTBLA, auch Kollegs bzw. Aufbaulehrgänge)	☐	☐
Fachhochschule, Pädagogische Akademie oder Universitätsstudium mit BA oder vergleichbarem Abschluss	☐	☐
Fachhochschule oder Universitätsstudium mit Mag., MA oder Dipl.-Ing.	☐	☐
Universitätsstudium mit Dr.	☐	☐
Ich weiß es nicht.	☐	☐

7d. Welchen Beruf übt deine Mutter aus? ➡ _____

7e. Wo arbeitet sie und was macht sie in ihrem Beruf normalerweise?

➡ _____

7f. Welchen Beruf übt dein Vater aus? ➡ _____

7g. Wo arbeitet er und was macht er in seinem Beruf normalerweise?

➡ _____

7h. Wie viele Bücher gibt es bei dir zu Hause?
Auf einen Meter Bücherregal passen ungefähr 40 Bücher.
➤ *Schätze bitte die Anzahl aller Bücher (egal in welcher Sprache sie sind),*
   aber zähle Zeitschriften, Zeitungen und deine Schulbücher <u>nicht</u> mit.

☐ zwischen 0 und 10 Bücher   ☐ zwischen 101 und 200 Bücher
☐ zwischen 11 und 50 Bücher   ☐ zwischen 201 und 500 Bücher
☐ zwischen 50 und 100 Bücher   ☐ mehr als 500 Bücher

DANKE für deine Mitarbeit!

Englisch in der Freizeit

universität wien

Sprachtagebuch

Liebe Schülerin, lieber Schüler!

In diesem Teil der Studie interessiert mich, was genau du diese Woche in deiner Freizeit mit Englisch machst und wie viel Zeit du in etwa mit Englisch verbringst.

Deshalb bitte ich dich, ab heute eine Woche lang ein kurzes **online Sprachtagebuch** auszufüllen. Das bedeutet, dass du dieselben Fragen an **7 Tagen** (z.B. von Dienstag bis Montag) beantworten sollst, also auch für Samstag und Sonntag.

Das Sprachtagebuch auszufüllen dauert ca. 5 min, du kannst es **am Handy, auf einem Tablet oder am Computer** ausfüllen. Bitte nimm dir **am Ende des Tages** kurz Zeit dafür oder füll es **in der Früh für den vorigen Tag** aus. Wichtig ist nur, dass du dich noch gut erinnern kannst!

Du wirst in den Fragen gebeten zu schätzen, wie viel Zeit du mit Aktivitäten verbracht hast – denk dabei bitte daran, **wie lange** du eine Aktivität **insgesamt pro Tag** gemacht hast.

Genauso wie der Fragebogen, sind deine Antworten auch im Sprachtagebuch **anonym** und werden **sicher** gespeichert und streng **vertraulich** behandelt.

Falls du Fragen hast, ruf mich an oder schreib mir einfach eine E-Mail!

Um das **Sprachtagebuch** zu **öffnen**, kannst du diesen Link verwenden: http://tinyurl.com/Sprachtagebuch

oder, wenn du es am Handy machst, diesen **QR Code** benutzen:

THANK YOU!

for your help!

Englisch in der Freizeit – Sprachtagebuch (*online*)

Liebe Schülerin, lieber Schüler!

Hier sind einige Fragen zu deinen Kontakten mit Englisch an einem bestimmten Tag. **Achtung**: Es geht hier nur um Kontakte mit Englisch in deiner Freizeit, die <u>nichts</u> mit der Schule zu tun hatten (also keine Hausübungen, Referatsvorbereitung, Lernen, Nachhilfe,…).

Damit ich deine Sprachtagebücher den anderen Teilen der Studie zuordnen kann, bitte ich dich hier deinen persönlichen Code einzufüllen:

Mein Code:	☐☐ 2. und 3. Buchstabe deines Vornamens
	☐ der letzte Buchstabe des Vornamens deiner Mutter
	☐☐ Geburtsmonat (z.B. 03 für März)

1. Für welchen Tag füllst du das Sprachtagebuch gerade aus?

Montag
Dienstag
Mittwoch
Donnerstag
Freitag
Samstag
Sonntag

2. Bitte gib auch das Datum an! (z.B. 18.04.2016)   **TEXTFELD**

3. Hast du am [*Tag der in Frage 1 angegeben wurde z.B. Montag*] in deiner Freizeit Englisch **gehört**? Was und wie lange?

	Nein	ca. 5 min	ca. 10 min	ca. 15min	ca. 30 min	ca. 45 min	ca. 1 Std	ca. 1,5 Std	ca. 2 Std	ca. 2,5 Std	ca. 3 Std oder mehr
Musik (auch Musikvideos)											
Radiosender (auch online)											
Filme (auch online)											
Serien (auch online)											
andere Fernsehsendungen (auch online)											
Videoclips											
in einem Spiel (z.B. PC, Konsole, Handy,..)											
etwas anderes											

3a. Was hast du am [Tag der in Frage 1 angegeben wurde z.B. Montag] noch auf Englisch gehört? [Frage wird nur angezeigt, wenn bei „etwas anderes" nicht „nein" ausgewählt wurde]

                                              TEXTFELD

3b. Hast du am [Tag der in Frage 1 angegeben wurde z.B. Montag] beim Ansehen der englischsprachigen Filme, Serien oder Sendungen Untertitel verwendet?
[*Frage wird nur angezeigt, wenn bei „Filme", „Serien" oder „Fernsehsendungen" nicht „nein" ausgewählt wurde*]
Ja
Nein

3c. In welcher Sprache waren diese Untertitel?
[*Frage wird nur angezeigt, wenn bei Frage 3a „Ja" ausgewählt wurde*]
Englisch
Deutsch
Andere Sprache: _____

4. Hast du am [*Tag der in Frage 1 angegeben wurde z.B. Montag*] in deiner Freizeit etwas auf Englisch **gelesen**? Was und wie lange?

	Nein	ca. 5 min	ca. 10 min	ca. 15 min	ca. 30 min	ca. 45 min	ca. 1 Std	ca 1,5 Std	ca. 2 Std	ca. 2,5 Std	ca. 3 Std oder mehr
Buch / E-book											
Zeitung / Zeitschrift											
Comics/Mangas											
Artikel im Internet											
Informationstexte im Internet (z.B. Wikipedia, Anleitungen)											
(Kurz)Geschichten im Internet											
Blog											
Songtexte											
E-Mails / Nachrichten in sozialen Medien											
Statusmeldungen / Kommentare in sozialen Netzwerken											
Einträge in einem Forum											
SMS/Whatsapp Nachrichten											
in einem Computerspiel											
etwas anderes											

4a. Was hast du am [*Tag der in Frage 1 angegeben wurde z.B. Montag*] noch auf Englisch gelesen?
[*Frage wird nur angezeigt, wenn bei „etwas anderes" nicht „nein" ausgewählt wurde*]

TEXTFELD

5. Was hast du am [*Tag der in Frage 1 angegeben wurde z.B. Montag*] in deiner Freizeit **gespielt** und wie lange?

	Nein	ca. 5 min	ca. 10 min	ca. 15 min	ca. 30 min	ca. 45 min	ca. 1 Std	ca. 1,5 Std	ca. 2 Std	ca. 2,5 Std	ca. 3 Std oder mehr
Computerspiele / Videospiele offline											
Computerspiele / Videospiele online											
Online Spiele mit anderen											
Spiele am Handy/ Tablet/Ipod oder anderen tragbaren Geräten											
etwas anderes											

5a. Was hast du am [*Tag der in Frage 1 angegeben wurde z.B. Montag*] noch auf Englisch gespielt? [*Frage wird nur angezeigt, wenn bei „etwas anderes" nicht „nein" ausgewählt wurde*]

TEXTFELD

6. Hast du am [*Tag der in Frage 1 angegeben wurde z.B. Montag*] etwas in deiner Freizeit auf Englisch **geschrieben**? Was und wie lange?

	Nein	ca. 5 min	ca. 10 min	ca. 15 min	ca. 30 min	ca. 45 min	ca. 1 Std	ca. 1,5 Std	ca. 2 Std	ca. 2,5 Std	ca. 3 Std oder mehr
Begriffe in einer Suchmachische (z.B. Google)											
E-Mails / Nachrichten in sozialen Medien											
Stautsmeldungen / Kommentar in sozialen Netzwerken											
SMS/Whatsapp Nachrichten											
in einem Chat (auch In-Game Chats bei Computerspielen)											
Einträge in einem Forum											
Blog											
Songtexte											
(Kurz)Geschichten (auch online)											
in einem Computerspiel											
Tagebuch											
etwas anderes											

6a. Was hast du am [*Tag der in Frage 1 angegeben wurde z.B. Montag*] noch auf Englisch geschrieben?
[*Frage wird nur angezeigt, wenn bei „etwas anderes" <u>nicht</u> „nein" ausgewählt wurde*]

TEXTFELD

7. Hast du am [*Tag der in Frage 1 angegeben wurde z.B. Montag*] in deiner Freizeit Englisch **gesprochen**? Was und wie lange?

	Nein	ca. 5 min	ca. 10 min	ca. 15 min	ca. 30 min	ca. 45 min	ca. 1 Std	ca. 1,5 Std	ca. 2 Std	ca. 2,5 Std	ca. 3 Std oder mehr
persönlich											
via Telefon											
via Skype oder ähnliche Internetdienste											
mit mir selbst											
auf Englisch gesungen (auch bei Liedern mitgesungen)											

7a. Mit wem hast du am [*Tag der in Frage 1 angegeben wurde z.B. Montag*] in deiner Freizeit Englisch gesprochen? [*Frage wird nur angezeigt, wenn bei „etwas anderes" <u>nicht</u> „nein" ausgewählt wurde*]

TEXTFELD

8. Hast du am [*Tag der in Frage 1 angegeben wurde z.B. Montag*] in deiner Freizeit noch anderen Sprachen verwendet?

Ich habe am [Tag der in Frage 1 angegeben wurde z.B. Montag] auch noch in einer anderen Sprache als Englisch oder Deutsch...

			Bitte gib alle Sprachen an!
	Ja	Nein	und zwar in...
etwas gelesen			
etwas gehört			
gesprochen			
geschrieben			

8a. Diesen [*Tag der in Frage 1 angegeben wurde z.B. Montag*] habe ich ... mit Englisch verbracht.
weniger Zeit als normalerweise
gleich viel Zeit wie immer
mehr Zeit als normalerweise

Herzlichen Dank für deine Mithilfe!
Bitte vergiss nicht, das Sprachtagebuch auch morgen wieder auszufüllen!

Produktiver Wortschatz: Wortassoziationen

Liebe Schülerin, lieber Schüler!

Damit ich diesen Teil den anderen Teilen der Studie zuordnen kann, bitte ich dich hier deinen persönlichen Code gut leserlich in BLOCKBUCHSTABEN einzufüllen:

Mein Code:	☐☐ 2. und 3. Buchstabe deines Vornamens
	☐ der letzte Buchstabe des Vornamens deiner Mutter
	☐☐ Geburtsmonat (z.B. 03 für März)

Hinweise zum Ausfüllen:

In dieser Überprüfung des Wortschatzes wirst du eine Liste mit 30 englischen Wörtern sehen.
*Schreibe neben jedes Wort, **die ersten anderen englischen Wörter, die dir dazu einfallen.***
*Schreibe so viele auf, wie dir einfallen, wenn möglich **vier Wörter** (neben jedem Wort sind vier leer Kästchen).*

Es ist egal, welche Verbindung es zwischen dem ersten Wort und deinen Wörtern gibt, schreib einfach auf Englisch auf, was dir dazu einfällt!
Zum Beispiel hat jemand für das Wort "animal" diese Wörter aufgeschrieben:

animal	elephant	farm	wild	feed

Denk nicht zu lange nach, du hast insgesamt nur <u>15 Minuten Zeit</u>!

Hast du noch Fragen?

1

Schreibe neben jedes Wort, die ersten anderen englischen Wörter, die dir dazu einfallen.

1. **attack**

2. **board**

3. **close**

4. **cloth**

5. **dig**

6. **dirty**

7. **disease**

8. **experience**

9. **fruit**

10. **furniture**

Schreibe neben jedes Wort, die ersten anderen englischen Wörter, die dir dazu einfallen.

11. **habit**				
12. **hold**				
13. **hope**				
14. **kick**				
15. **map**				
16. **obey**				
17. **pot**				
18. **potato**				
19. **real**				
20. **rest**				

Schreibe neben jedes Wort, die ersten anderen englischen Wörter, die dir dazu einfallen.

21. **rice**				
22. **science**				
23. **seat**				
24. **spell**				
25. **substance**				
26. **stupid**				
27. **television**				
28. **tooth**				
29. **trade**				
30. **window**				

Rezeptiver Wortschatz: Yes/No Aufgaben

Liebe Schülerin, lieber Schüler!

Damit ich diesen Teil den anderen Teilen der Studie zuordnen kann, bitte ich dich hier deinen persönlichen Code gut leserlich in BLOCKBUCHSTABEN einzufüllen:

Mein Code:	☐☐ 2. und 3. Buchstabe deines Vornamens
	☐ der letzte Buchstabe des Vornamens deiner Mutter
	☐☐ Geburtsmonat (z.B. 03 für März)

Hinweise zum Ausfüllen:

In dieser Überprüfung des Wortschatzes wirst du auf den nächsten Seiten Listen von Wörtern sehen.

*Bitte lies die Aufgaben **aufmerksam und genau** durch und kreuz für jedes Wort **Yes** oder **No** an:*

⇨ Wenn du weißt, was das Wort heißt, kreuze ☒ **Yes** an.
⇨ Wenn du nicht weißt, was das Wort heißt, kreuze ☒ **No** an.

Achtung:
Du musst vorsichtig sein, **zur Kontrolle sind Wörter eingebaut, die es im Englischen gar nicht gibt!**
Du bekommst Punkte abgezogen, wenn du so bei so einem erfundenen Wort **YES** anklickst!
Außerdem wird bei einigen Wörter später noch überprüft, ob du die Bedeutung tatsächlich kennst.

*Hier ist ein **Beispiel**:*

Block A	
which	☐ Yes ☐ No
industry	☐ Yes ☐ No

Block B	
siddy	☐ Yes ☐ No
doll	☐ Yes ☐ No

Hast du noch Fragen?

Ich weiß, was das Wort heißt:

Block A		
acute	☐ Yes	☐ No
podiast	☐ Yes	☐ No
malicious	☐ Yes	☐ No
fair	☐ Yes	☐ No
adjoin	☐ Yes	☐ No
makeshift	☐ Yes	☐ No
grudgingly	☐ Yes	☐ No
intimant	☐ Yes	☐ No
elphick	☐ Yes	☐ No
exemptation	☐ Yes	☐ No
rudge	☐ Yes	☐ No
misrequite	☐ Yes	☐ No
copper	☐ Yes	☐ No
todd	☐ Yes	☐ No
corn	☐ Yes	☐ No
greer	☐ Yes	☐ No
harridism	☐ Yes	☐ No
tranquil	☐ Yes	☐ No
leucan	☐ Yes	☐ No
costliness	☐ Yes	☐ No
sleek	☐ Yes	☐ No
dressy	☐ Yes	☐ No
elegance	☐ Yes	☐ No
carotic	☐ Yes	☐ No
insinuate	☐ Yes	☐ No
opie	☐ Yes	☐ No
breakwith	☐ Yes	☐ No
watler	☐ Yes	☐ No
spalding	☐ Yes	☐ No
strappery	☐ Yes	☐ No
concerned	☐ Yes	☐ No
pegler	☐ Yes	☐ No
analyse	☐ Yes	☐ No
shady	☐ Yes	☐ No
daintiness	☐ Yes	☐ No
fly	☐ Yes	☐ No
asbestial	☐ Yes	☐ No
engineer	☐ Yes	☐ No
keir	☐ Yes	☐ No
arbus	☐ Yes	☐ No

Block B		
bibby	☐ Yes	☐ No
liverick	☐ Yes	☐ No
flautism	☐ Yes	☐ No
greenaway	☐ Yes	☐ No
appreciate	☐ Yes	☐ No
wood	☐ Yes	☐ No
rumour	☐ Yes	☐ No
allaway	☐ Yes	☐ No
snape	☐ Yes	☐ No
bayonet	☐ Yes	☐ No
barmion	☐ Yes	☐ No
sedgebeer	☐ Yes	☐ No
boobier	☐ Yes	☐ No
sincere	☐ Yes	☐ No
undergraduate	☐ Yes	☐ No
application	☐ Yes	☐ No
detailoring	☐ Yes	☐ No
scamper	☐ Yes	☐ No
calves	☐ Yes	☐ No
stimulation	☐ Yes	☐ No
peebles	☐ Yes	☐ No
givewith	☐ Yes	☐ No
invest	☐ Yes	☐ No
integrality	☐ Yes	☐ No
surman	☐ Yes	☐ No
nonagrate	☐ Yes	☐ No
candish	☐ Yes	☐ No
harass	☐ Yes	☐ No
purchaser	☐ Yes	☐ No
redemption	☐ Yes	☐ No
decisively	☐ Yes	☐ No
vergial	☐ Yes	☐ No
convolition	☐ Yes	☐ No
displace	☐ Yes	☐ No
plagorate	☐ Yes	☐ No
alternate	☐ Yes	☐ No
ottery	☐ Yes	☐ No
frown	☐ Yes	☐ No
fresh	☐ Yes	☐ No
sample	☐ Yes	☐ No

Ich weiß, was das Wort heißt:

Block C		
advise	☐ Yes	☐ No
munch	☐ Yes	☐ No
reticence	☐ Yes	☐ No
honesty	☐ Yes	☐ No
couth	☐ Yes	☐ No
fearle	☐ Yes	☐ No
innoculism	☐ Yes	☐ No
expand	☐ Yes	☐ No
numb	☐ Yes	☐ No
ethical	☐ Yes	☐ No
gummer	☐ Yes	☐ No
resignate	☐ Yes	☐ No
cardination	☐ Yes	☐ No
pickard	☐ Yes	☐ No
verge	☐ Yes	☐ No
cranicle	☐ Yes	☐ No
gammonary	☐ Yes	☐ No
precocious	☐ Yes	☐ No
bought	☐ Yes	☐ No
ashill	☐ Yes	☐ No
swan	☐ Yes	☐ No
trill	☐ Yes	☐ No
derelict	☐ Yes	☐ No
ampled	☐ Yes	☐ No
acklon	☐ Yes	☐ No
ban	☐ Yes	☐ No
charlett	☐ Yes	☐ No
voluminary	☐ Yes	☐ No
rave	☐ Yes	☐ No
diversal	☐ Yes	☐ No
precious	☐ Yes	☐ No
postscript	☐ Yes	☐ No
castle	☐ Yes	☐ No
fastidious	☐ Yes	☐ No
safe	☐ Yes	☐ No
wookey	☐ Yes	☐ No
varney	☐ Yes	☐ No
maltass	☐ Yes	☐ No
murray	☐ Yes	☐ No
piccolotomy	☐ Yes	☐ No

Block D		
cartledge	☐ Yes	☐ No
obstinate	☐ Yes	☐ No
congulate	☐ Yes	☐ No
ostensibly	☐ Yes	☐ No
mastiphitis	☐ Yes	☐ No
refusal	☐ Yes	☐ No
steadily	☐ Yes	☐ No
scobie	☐ Yes	☐ No
prosaic	☐ Yes	☐ No
furrow	☐ Yes	☐ No
inhabitant	☐ Yes	☐ No
amiel	☐ Yes	☐ No
brind	☐ Yes	☐ No
elaborate	☐ Yes	☐ No
lone	☐ Yes	☐ No
morphew	☐ Yes	☐ No
guest	☐ Yes	☐ No
sprudd	☐ Yes	☐ No
incorpulent	☐ Yes	☐ No
haime	☐ Yes	☐ No
randle	☐ Yes	☐ No
albucolic	☐ Yes	☐ No
sanitary	☐ Yes	☐ No
menace	☐ Yes	☐ No
mourant	☐ Yes	☐ No
bluster	☐ Yes	☐ No
kearle	☐ Yes	☐ No
tomb	☐ Yes	☐ No
youde	☐ Yes	☐ No
asslam	☐ Yes	☐ No
scroll	☐ Yes	☐ No
lose	☐ Yes	☐ No
custom	☐ Yes	☐ No
tilt	☐ Yes	☐ No
seclunar	☐ Yes	☐ No
drab	☐ Yes	☐ No
gammage	☐ Yes	☐ No
powling	☐ Yes	☐ No
modesty	☐ Yes	☐ No
eventualise	☐ Yes	☐ No

Ich weiß, was das Wort heißt:

Block E		
practicate	☐ Yes	☐ No
motivise	☐ Yes	☐ No
pilbean	☐ Yes	☐ No
lucky	☐ Yes	☐ No
self-respect	☐ Yes	☐ No
raisin	☐ Yes	☐ No
compress	☐ Yes	☐ No
pocock	☐ Yes	☐ No
windle	☐ Yes	☐ No
churchlow	☐ Yes	☐ No
embarrassment	☐ Yes	☐ No
whereabouts	☐ Yes	☐ No
intuned	☐ Yes	☐ No
department	☐ Yes	☐ No
condick	☐ Yes	☐ No
shareholder	☐ Yes	☐ No
stephonise	☐ Yes	☐ No
interfere	☐ Yes	☐ No
scenery	☐ Yes	☐ No
mealing	☐ Yes	☐ No
discuss	☐ Yes	☐ No
tuber	☐ Yes	☐ No
thrift	☐ Yes	☐ No
orphan	☐ Yes	☐ No
squeak	☐ Yes	☐ No
fountain	☐ Yes	☐ No
hypodemical	☐ Yes	☐ No
possumate	☐ Yes	☐ No
cunnion	☐ Yes	☐ No
atribus	☐ Yes	☐ No
cordonise	☐ Yes	☐ No
pod	☐ Yes	☐ No
whaley	☐ Yes	☐ No
germ	☐ Yes	☐ No
mainwaring	☐ Yes	☐ No
floralate	☐ Yes	☐ No
cundy	☐ Yes	☐ No
remove	☐ Yes	☐ No
quote	☐ Yes	☐ No
cynist	☐ Yes	☐ No

Yes/No Übersetzungen

Liebe Schülerin, lieber Schüler!

Hier sind noch mal einige englische Wörter aus den Yes/No Aufgaben.
Bitte schreibe neben jedes englische Wort **eine deutsche Übersetzung**.

Wenn dir das deutsche Wort nicht einfällt, kannst du auch ein englisches Synonym oder eine Erklärung hinschreiben. (*In dieser Liste sind keine erfundenen Wörter dabei!*)

1. appreciate
2. shareholder
3. elegance
4. concerned
5. sincere
6. ethical
7. analyse
8. self-respect
9. tomb
10. frown
11. scenery
12. fresh
13. sample
14. engineer
15. fountain
16. orphan
17. embarrassment
18. honesty
19. quote
20. inhabitant

Du hast es geschafft!
Danke, dass du diese Wortschatzüberprüfung ausgefüllt hast!

Beantworte bitte zum Abschluss des heutigen Teils jetzt nur noch diese fünf kurzen Fragen:

⇨ *Wie fühlst du dich heute?*

☹	☹	😐	☺	☺
gar nicht gut	nicht so gut	okay	ziemlich gut	großartig

⇨ *Wie leicht war es heute für dich, dich auf die Aufgaben zu konzentrieren?*

☹	☹	😐	☺	☺
sehr schwierig	ziemlich schwierig	okay	ziemlich leicht	sehr leicht

⇨ *Hast du während der Tests Veränderungen in Aufmerksamkeit und Konzentration bemerkt?*

_____

⇨ *Wie hast du die Aufgabe mit den Wortassoziationen gefunden?*

☹	☹	😐	☺	☺
sehr schwierig	ziemlich schwierig	okay	ziemlich leicht	sehr leicht

⇨ *Wie einfach oder schwierig waren die **Yes/No Aufgaben** für dich?*

☹	☹	😐	☺	☺
sehr schwierig	ziemlich schwierig	okay	ziemlich leicht	sehr leicht

⇨ *Gab es Wörter, wo du dir nicht sicher warst, was du ankreuzen sollst? Wenn ja, wie viele circa?*

_____

⇨ *Hast du sonst noch irgendwelche Kommentare zum heutigen Teil?*

_____

Table A.1: Table of responses on the translation task accepted as correct and rejected as incorrect

Item No.	Item	Accepted translations		Accepted explanations		Not accepted
		Translations in bilingual dictionaries[1]	Student translations	Explanations based on monolingual dictionaries[2]	English/German explanations by students	
1	**appreciate**	(zu) schätzen (wissen), dankbar sein, Verständnis haben, sich bewusst sein, im Wert steigen	wertschätzen, anerkennen, akzeptieren und gern haben, annehmen	be grateful, value, understand, recognize worth, increase in value		gratulieren, gefallen, gutheißen, bevorzugen, sich an etw. erfreuen
2	**shareholder**	Teilhaber(in), Aktionär(in), Aktienbesitzer(in), Anteilseigner(in), Gesellschafter(in)		owner of shares, hold shares in company		Platzhalter
3	**elegance**	Eleganz	elegant	being elegant, stylish, graceful		
4	**concerned**	betroffen, besorgt	bekümmert, beunruhigt	worried, involved, affected by		bewusst, sicher sein, bezogen auf
5	**sincere**	ehrlich, aufrichtig		honest, free of deceit, genuine		Yours sincerely, Hochachtungsvoll, Mit freundlichen Grüßen
6	**ethical**	ethisch, moralisch		concerning morals and justice, good, correct, avoiding harm		Ethik, religiös
7	**analyse**	analysieren, untersuchen		examine in detail, identify and measure	take a closer look at	die Analyse
8	**self-respect**	Selbstachtung	Selbstrespekt, selbst - respektierend, Selbstwertgefühl	confidence in oneself	self-esteem Respekt für/vor sich selbst, sich selbst akzeptieren	Selbstrespect, persönliche Wertschätzung
9	**tomb**	Grab, Gruft, Grabkammer		vault, hole for burying the dead, burial place	a dead person lays in a tomb	Sarg, Höhle, Schatzkammer
10	**frown**	die Stirn runzeln, missbilligen	Grimasse, Gesichtsausdruck	furrow one's brow, disapprove	to raise your eyebrow, when someone's worried or not in a good mood Gesichtsausdruck wenn jemandem etwas nicht gefällt, Gegenteil von Lächeln	ängstlich, erstaunt sein (negativ), Stirn

11	**scenery**	Landschaft, Bühnenbild	Umgebung, Kulisse	landscape, painted background on stage, surroundings		die Szenerie, Szenario, Ortschaft, Aussicht
12	**fresh**	neu, ungebraucht, frisch, ausgeruht, kräftig (nur für Wind)		new, newly made, recent, unused, not stale, not processed, clean and pleasant, cool and windy, not tired, refreshing		
13	**sample**	Probe, Muster, Querschnitt, Stichprobe, ausprobieren, kosten, probieren, Probe entnehmen	Beispiel, Beispiel/Auszug, Testgröße eines Produkts	small amount, representative group, take a sample, try the quality, record or extract music digitally		Stück
14	**engineer**	Ingenieur(in), Techniker(in), Maschinist(in), Lokführer(in), konstruieren, etw bauen, arrangieren, aushecken	Techniker	someone who designs/constructs/works with machinery, buildings, bridges, engine driver, to arrange, skilfully plan		Arbeiter, Motor, Mechaniker, Engineer
15	**fountain**	Brunnen, Springbrunnen, Wasserstrahl, Quelle, spritzen,	Fontäne, Wasserspeier	ornamental jet/structure with water, jet of water, spray, source		Fontanne (spelling), Wasserfall
16	**orphan**	Waise, Waisenkind, verwaist		a child whose parents are dead		
17	**embarrass-ment**	Peinlichkeit, Verlegenheit, peinlich sein, Beschämung, Scham	Entblößung, bloß stellen, Demütigung, Blamage	self-consciousness, shame, awkwardness		embarrassed, when you do something wrong you feel that
18	**honesty**	Ehrlichkeit	Aufrichtigkeit, Wahrheit	quality of being honest		Ernsthaftigkeit
19	**quote**	Zitat, Kostenvoranschlag, Angebot, zitieren, ansetzen (Preis), anführen	Spruch, Aussage, Anführungszeichen	quotation, estimate, repeat a statement, mention, refer to	like a sentence, a famous sentence a person says	Quote, Gedicht
20	**inhabitant**	Einwohner(in), Bewohner(in)	bewohnen	person or animal who lives (permanently) in a place		Lebensraum, Ureinwohner, Inländer, unbewohnbar

[1] The bilingual dictionaries used were Cambridge online dictionary English-German (Cambridge University Press 2018) and PONS online dictionary English-German (PONS GmbH 2018).

[2] The monolingual dictionaries used were Cambridge online dictionary English (Cambridge University Press 2018) and Oxford living dictionaries English (Oxford University Press 2018).

<div align="center">

Leitfaden
zur Verwendung in den Fokusgruppeninterviews

</div>

<u>Einleitung</u>

Vielen Dank nochmals, dass ihr euch Zeit nehmt, heute gemeinsam noch ein paar Fragen zum Forschungsprojekt „Englisch in der Freizeit" zu diskutieren. Ich habe wie versprochen Pizza/Snacks und Getränke für euch vorbereitet, bitte bedient euch!

In dieser Gruppendiskussion geht es darum, was <u>ihr</u> denkt, es gibt hier keine richtigen oder falschen Antworten. Und wie ihr seht, steht hier ein Aufnahmegerät. Wie schon gesagt, ist es deswegen da, weil ich es nie schaffen würde, alle eure Gedanken so schnell mitzuschreiben. Ich werde mir nur ein paar Notizen machen können, aber die Aufnahme wird abgetippt und dann wird der anonymisierte Text weiter verwendet und niemand außer mir und dem Studenten, der mir hilft das alles abzutippen, wird diese Aufnahme hören.

Deshalb möchte ich euch auch bitten, euch selbst einen Decknamen aussuchen, den wir dann statt eurem Namen verwenden werden, damit niemand weiß wer was gesagt hat und es wird auch niemand wissen, dass es an dieser Schule war.

*Nur ganz kurz bevor wir starten noch ein paar wenige **Diskussionsregeln**:*

⇨ *Wir hören einander zu und lassen andere ausreden (sonst wird es zum Aufschreiben ein bisschen schwierig).*
⇨ *Es ist völlig okay, wenn ihr einander antwortet und auch ihr könnt einander natürlich Fragen stellen.*

<u>Zur Bedeutung von Englisch für Jugendliche</u>

Na dann, fangen wir gleich mit der ersten Frage an! Ich würde euch bitten, dass dazu jeder von euch kurz etwas sagt.
Wie wichtig ist Englisch in eurem Alltagsleben im Vergleich zu anderen Sprachen?

Gibt es eurer Meinung nach etwas, das man auf Englisch besser machen kann als in anderen Sprachen?

Könntet ihr euch vorstellen, wie es wäre ein Jugendlicher/eine Jugendliche in Österreich zu sein und kein Englisch zu können?
⇨ Was könnte man eurer Meinung nach ohne Englisch nicht tun?
⇨ Wo würde es euch am meisten fehlen?

<u>Zum Umgang mit außerschulischem Englisch</u>

Ihr habt ja auch einen Fragebogen ausgefüllt, insgesamt haben diesen 216 Wiener SchülerInnen wie ihr gemacht. **Ich habe hier die Ergebnisse der Umfrage zu den beliebtesten englischen Freizeitaktivitäten** [*Ergebnisse herzeigen/auf Tisch verteilen*].

Wie ihr seht, sind englische Musik hören und englische Videoclips ansehen, die Aktivitäten die Jugendliche am häufigsten auf Englisch machen. Blau steht nämlich für fast jeden Tag, orange für ein paar Mal in der Woche und grau für weniger oft. Außerdem hat mehr als die Hälfte der Jugendlichen gesagt, dass sie zumindest ein paar Mal pro Woche Filme und Serien auf Englisch

ansehen und dass sie Englisch in sozialen Netzwerken lesen. Mehr als die Hälfte verwendet mehrmals pro Woche englische Wörter in anderen Sprachen, für Suchmaschinen wie Google und auch für Apps.

Was sagt ihr dazu?

⇨ Überrascht euch dieses Ergebnis?

⇨ Warum glaubt ihr, dass genau diese Freizeitaktivitäten am öftesten auf Englisch gemacht werden?

Die Studie hat auch ergeben, dass Wiener SchülerInnen im Durchschnitt etwas mehr als 4 Stunden pro Tag mit Englisch verbringen. Was sagt ihr zu diesem Ergebnis?

⇨ Überrascht euch das?

⇨ Trifft diese Schätzung auch auf euch zu?

⇨ Was glaubt ihr, dass es tatsächlich bedeutet, wenn Leute im Sprachtagebuch angegeben haben, dass sie 4 Stunden oder mehr mit Englisch verbracht haben?

Könnt ihr irgendwie beschreiben, welche Arten von Englisch ihr außerhalb der Schule hört oder lest?

Zur Bedeutung von außerschulischem Englisch für den Sprachlernprozess

Denkt ihr, dass euch die außerschulische Beschäftigung mit Englisch beim Englischlernen hilft?

⇨ Warum (nicht)?

Was kann man eurer Meinung nach außerhalb der Schule lernen?

Könnt ihr Beispiele für etwas geben, dass ihr durch die Beschäftigung mit Englisch außerhalb der Schule gelernt habt?

Von welchen englischen Freizeitaktivitäten kann man etwas lernen?

⇨ Ihr seht ja die beliebtesten englischem Freizeitaktivitäten vor euch [*Bilder EE activities*], gibt es da eurer Meinung nach Unterschiede, was oder wieviel man davon lernen kann? [*nachhaken*]

⇨ Gibt es Aktivitäten, die ihr nicht oder nicht so oft macht, von denen man aber vielleicht etwas lernen könnte?

⇨ Kommuniziert ihr auch auf Englisch in eurer Freizeit?
z.B. auch mit englischen Native speakers?

Gibt es irgendwelche Ratschläge, die ihr anderen Jugendlichen geben würdet, damit sie von Englisch außerhalb der Schule profitieren können?

Außerschulisches Englisch und der Wortschatzerwerb

[*Bei Wortschatzbeispielen einhaken*]

Ihr habt als Beispiel für außerschulisches Lernen auch Wörter und Vokabel genannt und ihr wisst ja, dass es in diesem Forschungsprojekt auch um den Wortschatz geht. Im Fragebogen gab es auch eine Frage zu was Jugendliche mit neuen englischen Wörtern in der Freizeit machen.

Hier seht ihr die Ergebnisse [*Ergebnisse Vokabellernstrategien zeigen*], also mehr als die Hälfte von euch macht nicht irgendetwas, wenn ihnen ein neues englisches

Wort begegnet, weil grau steht für fast nie und wir sehen ganz oben, dass 59,9% fast nie nichts machen. Was Wiener Jugendliche machen ist unterschiedlich, viele erschließen sich die Bedeutung aus dem Zusammenhang oder schauen im Wörterbuch nach, oder manche denken auch an andere Sprachen. Was weniger oft passiert ist, dass Jugendliche jemand fragen, oder das Wort in Teile zerlegen oder sich überlege was für eine Wortart das ist.

Könnt ihr mir mehr darüber erzählen, wie ist das bei euch ist?
⇨ Könnt ihr Beispiele geben?
⇨ Ich find das spannend aus dem Zusammenhang erschließen: was heißt das dann?
⇨ Warum glaubt ihr ist das die häufigste Strategie?
⇨ Macht ihr noch etwas, das da nicht dabei ist?

Von welchen englische Freizeitaktivitäten, kann man besonders gut Vokabel lernen?
⇨ Könnt ihr konkrete Beispiele geben, welche Wörter ihr da lernt ?
⇨ Welche Wörter kommen euch da unter?
⇨ Wo verwendet ihr diese Wörter dann?

Außerschulisches Englisch und die Schule

In dieser Studie geht es um die Freizeit, also euren außerschulischen Kontakt mit Englisch. Gleichzeitig habt ihr aber auch (schon viele Jahre) Englischunterricht in der Schule. Beeinflussen sich das außerschulische Englisch und der Englischunterricht gegenseitig?

(Also haben eure englischen Freizeitaktivitäten einen Einfluss auf den Englischunterricht? Und umgekehrt?)
⇨ Könnt ihr Bespiele geben, wie sich außerschulisches Englisch und Englisch inder Schule gegenseitig beeinflussen?
⇨ Das was ihr außerhalb der Schule macht, spielt das eine Rolle im Englischunterricht?
⇨ Sollen Englisch in und außerhalb der Schule überhaupt in Kontakt kommen? /
Sollen eure englischen Freizeitbeschäftigungen im Unterricht eine Rolle spielen?
⇨ Würdet ihr euch wünschen, dass Lehrer mehr achten darauf, was ihr in der Freizeit macht?

Denkt ihr, dass dieses Projekt eure Bewusstheit/Aufmerksamkeit über Englisch außerhalb der Schule verändert hat?
⇨ Könnt ihr beschreiben, wie es sich verändert hat?
⇨ Habt ihr z.B. durch das Sprachtagebuch mehr auf Englisch in eurer Umgebung geachtet?

Abschluss

Möchtet ihr mir sonst noch irgendetwas zum Thema Englisch in der Freizeit oder generell sagen?

Vielen Dank für eure Hilfe. Ich hoffe, ihr habt diese Diskussion genauso interessant gefunden wie ich, ich habe sehr interessante Dinge gelernt!

Musik hören
9,9%
10,9%
79,2%
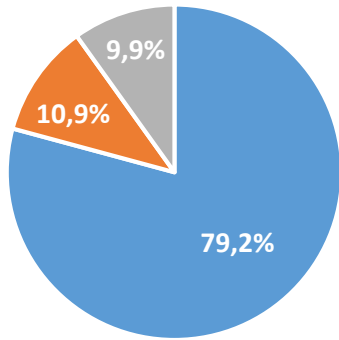
Videoclips ansehen
20,3%
25,7%
54,0%

Filme ansehen
29,6%
40,8%
29,6%
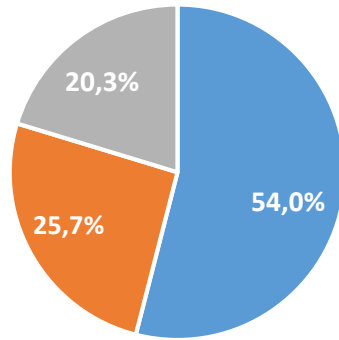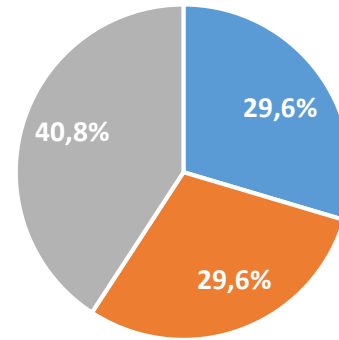
Serien ansehen
28,7%
45,6%
25,7%
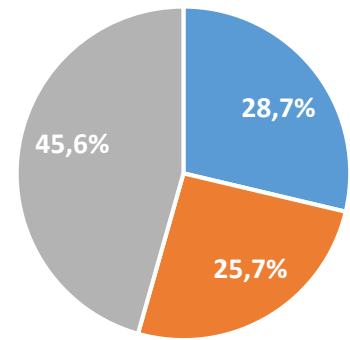
Die beliebtesten englischen Freizeitaktivitäten bei Wiener SchülerInnen

in sozialen Medien lesen
23,3%
20,3%
56,4%
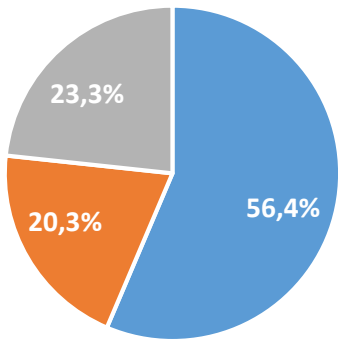
englische Wörter in anderen Sprachen verwenden
37,1%
18,3%
44,6%
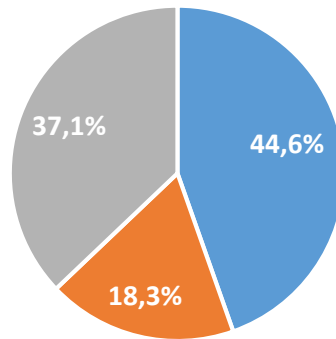
Suchmaschinen verwenden
29,7%
49,5%
20,8%
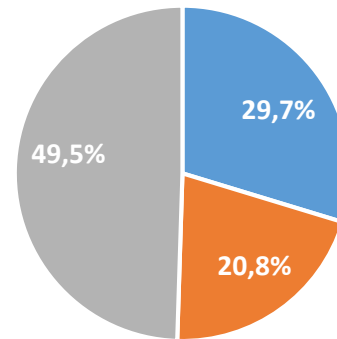
Apps nutzen
32,7%
48,5%
18,8%
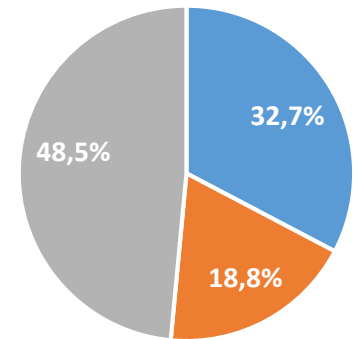
■ (fast) jeden Tag   ■ ein paar Mal pro Woche   ■ weniger oft
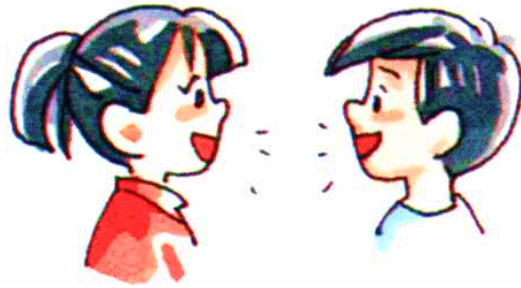
Freizeitaktivitäten
auf Englisch

Was Wiener SchülerInnen machen, wenn ihnen ein neues englisches Wort in der Freizeit begegnet

	sehr oft	oft	selten	(fast) nie
nichts	4,3%	8,0%	27,8%	59,9%
jemand fragen	13,2%	27,0%	30,7%	29,1%
Wörterbuch	47,6%	27,8%	15,5%	9,1%
aus dem Zusammenhang in Texten erschließen	45,5%	43,9%	7,9%	2,9%
aus dem Zusammenhang in Filmen/Serien erschließen	40,7%	42,9%	8,5%	7,9%
an andere Sprachen denken	29,6%	40,7%	16,9%	12,7%
in Wortteile zerlegen	13,2%	38,6%	28,6%	19,6%
Wortart überlegen	5,8%	15,3%	34,9%	43,9%

Legend: ■ sehr oft ■ oft ■ selten ■ (fast) nie
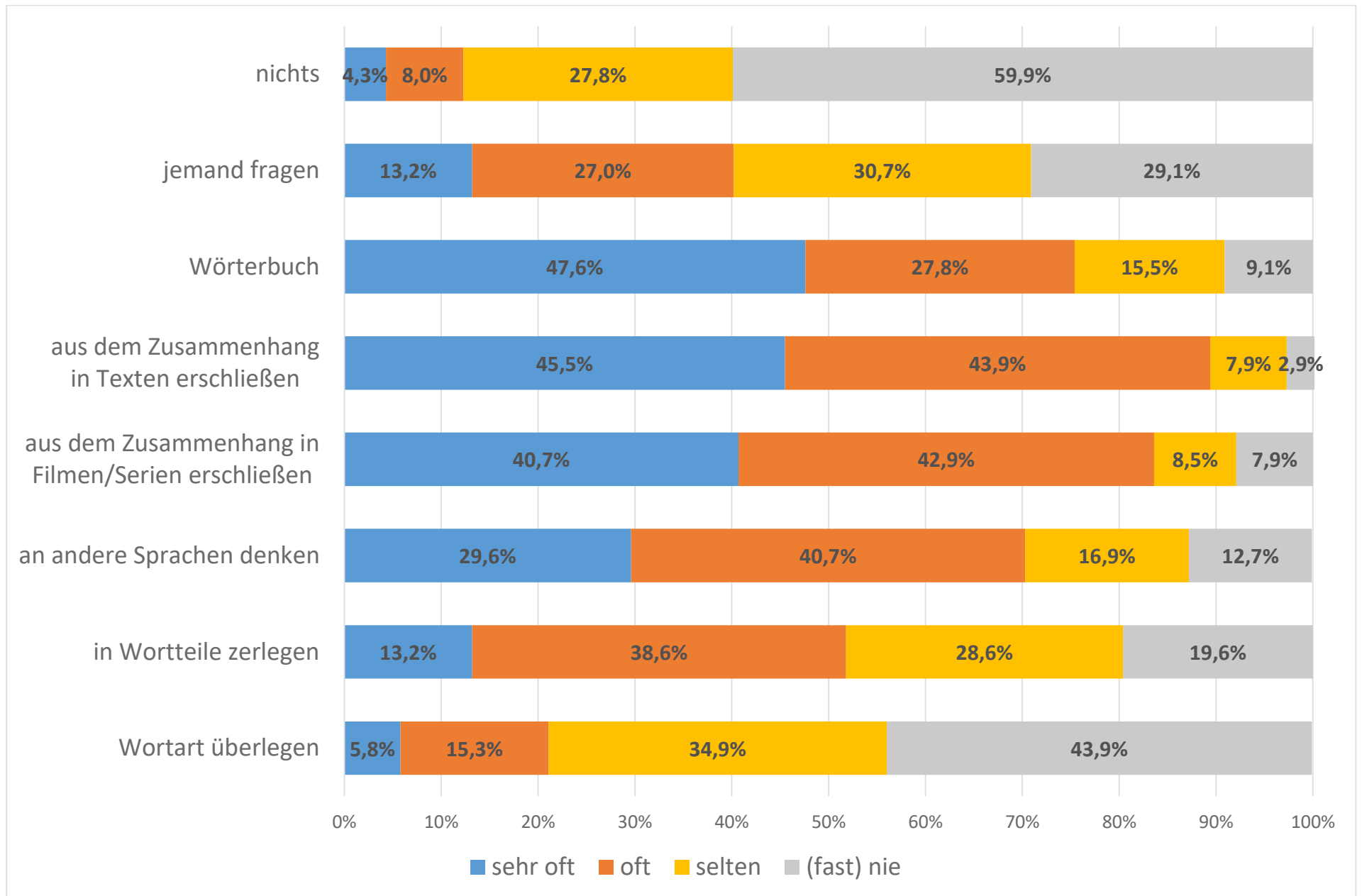
Transcription conventions

Example	Explanation
Interviewer:	Interviewer-ID
Pseudonym:	Speaker-ID for students = pseudonym
SS	Several students speaking at once
(Paul):	Statements was most probably made by the student given
SX-f, SX-m	Statements made by an unidentified female or male student
(.)	Pauses shorter than 1 second
(4)	Pauses longer than 1 second; time given in parentheses
Lukas: ist eigentlich alles auf <1> Englisch </1> Kira: <1> Ja bei </1> mir auch	Overlaps are numbered consecutively, all simultaneous overlaps have the same number
<ENGLISH> awesome </ ENGLISH > <SLOW> </SLOW> <WHISPERING> </WHISPERING> etc.	Statements expressed in a different language or in a particular way
@	Laughter: approximate number of syllables
Lia: Das (glaub ich auch)	Parentheses for statements that are not completely clear
John: das kommt vor allem bei xxxx vor	x for incomprehensible stretches of speech: approximate number of syllables
[someone enters the room] [children shouting outside]	Square brackets for additional information

Codebook used in the qualitative content analysis

Main category	Description	Categories	Description	Subcategories	Description
Significance of English in everyday life of young Austrians	Participants describe the significance of English in their everyday lives.	**Sociolinguistic roles of English in the lives of young Austrians**	Participants describe the sociolinguistic roles English fulfills in their everyday lives.	**English is not necessary in Austria**	Some participants argue that English is not necessary in Austria today.
				English has become normalized	English has become normal for some participants.
				Not knowing English would be horrible	Some participants argue that it would be horrific not to know English as an Austrian teenager today in a thought experiment.
		Reasons why English is important for young Austrians	Participants give reasons for why English is significant in their daily lives.	**It's the universal language ("Die Universalsprache")**	English is important because everybody speaks it around the world.
				Needed for leisure time activities	English is important for many of their leisure time activities to be able to watch original versions, take part in specific activities or understand online content.
				Importance for the future	English is important because you will need it in the future for further education or finding a job.
				For stays abroad/holidays	English is important to communicate on holidays or other stays abroad.
				Anglicisms in German	English is important because there are many English words used in Austrian German.
				Part of youth language	English is important because it has entered Austrian youth language.
				Finding information/ staying up-to-date	English is important to find information and staying up-to-date with the latest developments.
				(International) communication	English is important for international communication.

Appendix A

Description of EE	Participants describe their extramural English activities and reasons for doing them.	**Description of EE activities**	Participants describe what they do with English outside school in their leisure time.	**Participants' current EE activities**	Participants describe (some of) their current EE activities.
				Time spent on EE	Participants describe how much time they spend on various EE activities.
		Reasons for using English	Participants describe the reasons for using English in their everyday life with a focus on extramural activities.	**Many things just are in English**	English is the main language in certain environments, especially online, so that many things are mainly available in English.
				Original version	Many media are originally produced in English and these original versions are described as better than German translations.
				Wider pool of information	English is seen as granting access to a wider pool of information (especially in online environments).
				Friends or family abroad (who don't speak German)	English is needed for communication with friends or family who live abroad and commonly do not speak German.
				International communication (while gaming)	English is needed for online communication when engaging in certain activities, especially multiplayer games.
				Availability and being up-to-date	It takes time until German translations or dubbed versions are available, thus one has to use English to stay up to date.
				Pop culture is influenced by English-speaking countries	English is important because pop culture is strongly influenced by English-speaking countries and especially the US.
				English is cool or beautiful	English is described as being beautiful or cool, especially in relation to German.
				Easier to express oneself in English (more choice)	One can express oneself better in English as there is more (lexical) choice and it is difficult to translate the same content into German.
				Better content	English media provide better content. [maybe cut]
				Not having to switch language	When doing an activitiy in English and doign another activity, it is easier to stay in English than to switch languages.
				Intentional learning	Doing activities in English with the purpose of learning from them.

				Used as a secret language	English is used as as secret langauge in specific situations where othes do not speak it.
Learning from EE	Participants describe different aspects of learning from extramural English and/or evaluate EE's potential for language learning.	**What can be learned from EE**	Participants describe what they think can be learned from EE.	**Vocabulary**	Vocabulary can be learned from EE.
				Pronunciation	Pronunciation including stress and intonation can be learned from EE, one is also confronted with different pronunication variants.
				Idioms and Phrases	Idioms and phrases can be learned from EE.
				Casual/colloquial English or slang	Casual or colloquial English ("Umgangssprache") and slang expressions can be learned from EE.
				More natural English	Engaging with EE helps to develop a sense for more natural English, which can be related to casual rather than formal English but not necessarily.
				Procedural knowledge (rather than declarative)	Procedural knowledge of how English sounds or is used - "a better feeling" for the language can be learned from EE.
				Speaking (freely)	EE helps with developing (free) speaking skills.
				Better comprehension skills	EE helps with developing comprehension skills.
				Spelling	Spelling can be learned from EE.
				Grammar	Grammatical structures can be learned from EE.
		Evaluation of learning from EE	Participants evaluate their learning gains from EE - currently and in the past.	**Positive evaluation of learning from EE**	Learning from EE is evaluated positively.
				Mixed evaluation of learning from EE	Mixed views are expressed about learning from EE or participants change their opinion.
				Negative evaluation of learning from EE	Learning from EE is evaluated negatively.
		Evaluation of learning	Participants evaluate the potential of learning for	**Positive evaluation of learning potential**	The learning potential is evaluated positively.

		potential of EE activities	different EE activities, this category differs from "Evaluation of learning" because they do not necessarily engage in these activities.	**Mixed evaluation of learning potential**	Mixed views are expressed or participants disagreee about the learning potential of activities.
				Negative evaluation of learning potential	The learning potential is evaluated negatively.
		What helps learning from EE	Participants describe positive learning experiences from EE and explicitly or implicitly reveal factors that help learning from EE in their opinion.	**Motivation and own interest**	Since EE activities are done voluntarily and because of participants' interests, the motivation to understand the content and therefore look up or think about language is higher.
				Repetition	Repeated encounters with the same language structures help to learn from EE activities.
				Familiarization effect	EE helps to become familiar with ways English is used in certain media or with different kinds of English over time ("Gewöhnugnseffekt").
				English native speakers	English can be learned better from native speakers, either through direct contact or 'authentic' input.
				(Having to) use English actively	Using English actively in their free time through speaking or writing is helpful - especially if you have to use English as others do not speak German (though they are not necessarily native speakers of English).
				Linking English to audiovisual content	Integrating visual images, audio input and information on the plot helps to learn and remember language structures.
				Seeing language in written form	Seeing language in written form helps to learn and remember language structures.
				Being corrected by others	Being corrected by others when making mistakes helps to learn, both in physical and online setting.
				Using or triggering previous knowledge	Already known langauge structures are recycled and thus consolidated by encountering them in EE activities.
				Collaborative solving of a language problem	Collaboration to solve a language problem or question in online settings can support learning from EE.

				Learning wrong things from people who speak "bad English"	Engaging with people, especially non-native speakers, who speak 'bad English' holds the danger of 'unlearning English' or picking up wrong structures.
		Problems with learning from EE	Participants describe problems they have experienced with learning from EE and factors that hinder learning from EE in their opinion.	In songs language is different	In songs language is used differently in terms of word order or pronunciation which is not helpful for learning.
				Learning incorrect grammar or spelling	In EE input one encounters wrong spellings and (prescriptively) inaccurate grammatical constructions.
				Making wrong inferences	There is a danger of making wrong inferences about the meaning of an unknown lexical item or language structure if one does not check.
				Not encountering language structure often enough	Words one does not know yet do not come up often enough to learn outside school.
				Not learning anything new due to repetition	In some activities there is nothing new to learn once one has reached a certain level.
				English sometimes too difficult	Sometimes the level of English is too high and there is too much technical language to understand.
		Examples of learning from EE	Participants describe learning experiences that involve EE activities.		
		Learning vocabulary from EE	Participants describe experiences of learning vocabulary from EE.	Use of strategies	Participants describe their use of vocabulary learning strategies and give reasons why some are used more than others.
				Which words are looked up	Participants describe which unknown words they look up or use other strategies for.
				Remembering new words	Particiapnts describe what helps them to remember new words they encounter in their EE activities.
The relationship between in- and out-of-	Participants describe the relation between in- and out-of-	**Links between out-of-school English and English lessons**	Participants describe different kinds of links between English inside and outside of school.	No link	In participants' perception there is no link between English in and out of school.
				Mutual influence	English in and out of school mutually influence each other.

school English	school language learning drawing on their experiences.			**EE helps with English at school**	Engaging with English outside school helps with or has positive effects on English at school.
				EE creates problems with English at school	Engaging with English outside school creates problems in English lessons at school.
				Basics are learned at school then more from EE	You need to (be forced to) learn the basics at school but once you have reached a certain level (definitely in upper secondary) you learn more outside school.
				Learn more outside school than in school (now)	They have learned more outside school or they would never know as much if the had English just at school.
Evaluation of quantitative results	Participants evaluate selected results from the quantitative strand of the study.	**Evaluation of most frequent EE activities**	Participants state their opinion on the most frequent EE activities according to the survey.		
		Evaluation of time spent with EE	Participants state their opinion on mean time spent with EE according to the EEOLD.		
		Evaluation of VLS	Participants state their opinion on selected results of the quantitative strand.		
Types of English	Participants describe their contact with as well as their opinions about different types of English both inside and outside the school context.	**Varieties of English**	Participants describe their contact with and opinions on regional varieties of English.		
		Register differences	Participants describe differences in register and other aspects in the types of English they encounter outside school.		
English and other languages	Participants talk about the relations of English and other languages in their environment.	**English is more important than other foreign languages**	English is described as more important than other languages participants learn at school, and for some as equally though not more important than their L1s.		
		English is evaluated as	Similar to "*Description of EE|Reasons for using*		

		'better' than German	*English|English is cool or beautiful*" - English is seen as having better aesthetic qualities or being cooler than German.	
		The influence of English on German	The influence of English on German (and other languages), especially the use of Lenglish words in other languages, is evaluated both positively and negatively.	

Appendix B

Appendix B

Table B.1: A comparison of self-assessment sum scores and binned self-assessment scores

Self-assessment sum score	frequency	%	Binned self-assessment score	frequency	%
4	2	1.06	A2	7	3.72
5	5	2.66			
6	11	5.85	A2/B1	11	5.85
7	14	7.45	B1	80	42.55
8	29	15.43			
9	37	19.68			
10	30	15.96	B1/B2	30	15.96
11	34	18.09	B2	60	31.91
12	26	13.83			

N=188

Details on the binning procedure:

For each student four self-assessment scores are available corresponding to a rating of proficiency for each of the four skills on a scale using the three CEFR levels A2, B1 and B2. During questionnaire coding these level were assigned a numerical code, so that 1 corresponds to A2, 2 to B1 and 3 to B2. Thus, these ratings can be aggregated into one sum score (see left part of Table 6a.1 above), which is however hard to interpret in terms of language proficiency. Therefore, the sum score was cut into levels representing proficiency levels of the CEFR in a process called binning. The reasoning behind the binning transformation is as follows: A participant who rated themselves at level A2 for all four skills has a self-assessment sum score of 4 (=4 × 1). Pursuing this logic a sum score of 5 means a student rated three skills at A2 (= 3 × 1) and one skill at B1 (=1 × 2), but a sum score of 6 can either mean that a student rated three skills at A2 (= 3 × 1) and one skill at B2 (=1 × 3) or, and that is perhaps more likely, two skills at A2 (= 2 × 1) and two skills at B1 (=2 × 2). Therefore, a sum score of 6 has to be interpreted as in between levels A2 and B1, which is indicated in the right part of Table 6a.1 above. A sum score between 7 and 9 then indicates that a participant rated themselves mostly at B1 level (or two skills at A2 and two skills at B2, but this seems improbable), a score of 10 lies in between B1 and B2 and a score of 11 or 12 means that at least three skills were rated at level B2.

Table B.2: Frequency of general leisure times pursuits

	almost never	%	a few times a year	%	a few times a month	%	a few times a week	%	almost daily	%	*N*
listening to music	3	1.59	0	0	5	2.65	17	8.99	164	86.77	189
listening to the radio	67	35.64	13	6.91	46	24.47	30	15.96	32	17.02	188
watching TV	20	10.64	17	9.04	35	18.62	69	36.7	47	25	188
watching films or series on DVD	53	28.04	56	29.63	48	25.4	27	14.29	5	2.65	189
watching films or series on the internet	16	8.47	10	5.29	45	23.81	55	29.1	63	33.33	189
using social network sites (e.g. Facebook, Instagram,...)	10	5.32	0	0	4	2.13	21	11.17	153	81.38	188
watching video clips on the internet	2	1.06	1	0.53	7	3.72	37	19.68	141	75	188
playing games on your own (on a computer, console or online)	68	35.98	26	13.76	44	23.28	24	12.7	27	14.29	189
playing games online with others (e.g. Multiplayer Online Games)	89	47.59	21	11.23	26	13.9	28	14.97	23	12.3	187
playing games on a phone or tablet	34	18.28	31	16.67	57	30.65	33	17.74	31	16.67	186
reading newspapers or magazines (also online)	25	13.44	20	10.75	55	29.57	61	32.8	25	13.44	186
reading books or e-books	48	25.4	52	27.51	42	22.22	29	15.34	18	9.52	189
listening to audiobooks	162	86.17	13	6.91	8	4.26	4	2.13	1	0.53	188
meeting friends	4	2.14	5	2.67	20	10.7	82	43.85	76	40.64	187
doing sports	9	4.81	5	2.67	36	19.25	98	52.41	39	20.86	187
making music (e.g. singing or playing an instrument)	90	48.13	12	6.42	18	9.63	34	18.18	33	17.65	187
going to concerts	100	53.19	65	34.57	19	10.11	4	2.13	0	0	188
going to the cinema	13	6.95	107	57.22	60	32.09	7	3.74	0	0	187
going to the theatre	114	60.64	60	31.91	14	7.45	0	0	0	0	188
other activities	0	0	1	1.67	17	28.33	17	28.33	25	41.67	60

Table B.3: Gender differences for general leisure time activities

	N	Mdn fe-male	95% CI [1]	Mdn male	95% CI	W	p	r	95% CI
listening to music	188	5		5		4068	.275	-.08	[-.22, .07]
listening to the radio	187	3	[3, 3]	3	[1, 3]	4844	.093	-.12	[-.26, .02]
watching TV	187	4	[4, 4]	4	[3, 4]	4896	.074	-.13	[-.27, .01]
watching films or series on DVD	188	2	[2, 3]	2	[1, 2]	4965.5	.064	-.14	[-.27, .01]
watching films or series on the internet	188	4	[4, 4]	4	[3, 4]	4557	.478	-.05	[-.19, .09]
using social network sites (e.g. Facebook, Instagram,…)	187	5		5	[5, 5]	5239.5	<.001**	-.29	[-.41, -.15]
watching video clips on the internet	187	5		5		3830.5	.116	-.11	[-.26, .03]
playing games on your own (on a computer, console or online)	188	1	[1, 2]	4	[3, 4]	1624.5	<.001**	-.55	[-.64, -.44]
playing games online with others (e.g. Multiplayer Online Games)	186	1		4	[3, 4]	883.5	<.001**	-.72	[-.78, -.64]
playing games on a phone or tablet	186	3	[2, 3]	4	[3, 4]	2582	<.001**	-.34	[-.46, -.02]
reading newspapers or magazines (also online)	185	3	[3, 3.5]	4	[3, 4]	3579.5	.096	-.12	[-.26, .02]
reading books or e-books	188	3	[2, 3]	2	[2, 2]	5638	<.001**	-.27	[-.40, -.13]
listening to audiobooks	187	1		1		4455	.354	-.07	[-.21, .08]
meeting friends	186	4	[4, 5]	4	[4, 5]	4293.5	.842	-.01	[-.16, .13]
doing sports	186	4		4		3597	.064	-.14	[-.28, .01]
making music (e.g. singing or playing an instrument)	186	3	[2, 4]	1	[1, 2]	5137	.006**	-.20	[-.34, -.06]
going to concerts	187	2	[1, 2]	1	[1, 1]	5098.5	.011*	-.19	[-.14, .14]
going to the cinema	186	2	[2, 2]	2	[2, 2]	4261	.841	-.01	[-.16, .13]
going to the theatre	187	1	[1, 2]	1	[1, 1]	4639.5	.217	-.09	[-.23, .05]

p < .05, ** *p* < .01

[1] The empty cells in the columns presenting the 95% confidence intervals for the median in the female and male group could not be calculated due to computational issues: an error that all values in the bootstrapping sample corresponded to the corresponding median and that therefore CIs could not be computed was reported.

Table B.4: Analysis of qualitative EEQ data on most frequently used websites

Websites	Frequency according to response number					Total
	1g response 1	1g response 2	1g response3	1g response 4	1g response 5	
9Gag[3]	2	2	2	1	2	9
Amazon	2	2	0	4	2	10
ask.fm	0	0	1	4	2	7
Burning series	8	8	6	4	2	28
Facebook	5	16	8	8	4	41
Google	24	17	9	6	6	62
Google Translate[4]	1	1	3	0	1	6
Instagram	24	27	17	12	6	86
LeoDict[4]	0	1	2	0	0	3
Netflix	2	2	3	5	4	16
orf.at[2]	0	2	2	0	0	4
Pinterest	0	0	2	0	0	2
Reddit[2]	0	0	2	2	0	4
Snapchat	2	17	6	6	3	34
Soundcloud[1]	0	2	3	0	1	6
Sportify[1]	0	1	1	0	0	2
Tumblr	0	2	4	1	2	9
Twitter	3	4	6	4	0	17
Wattpad	0	1	0	1	1	3
WebUntis	1	0	1	1	0	3
WhatsApp	16	4	11	1	5	37
Wikipedia	3	9	10	9	4	35
YouTube	76	39	11	11	5	142
Zalando	1	0	1	1	0	3
Categories						
adult site	1	2	3	0	0	6
advice website	0	0	1	0	1	2
anime	2	0	0	1	0	3
browser	0	0	1	1	0	2
comics/mangas	0	2	0	0	0	2
cooking	0	0	0	1	0	1
dictionary/translator[4]	0	3	0	2	1	6
drawing	1	0	0	0	1	2
e-mail	0	2	0	2	1	5
events	0	0	0	0	1	1
Fan site	0	0	2	1	0	3
fashion	0	2	0	0	1	3
fun/satire[3]	0	0	1	0	2	3
gaming[5]	0	3	4	5	3	15
gaming/shopping[5]	0	0	0	1	0	1

images	0	0	0	0	1	1
IT	0	0	1	0	0	1
magazine	0	0	0	1	0	1
motor	0	0	1	1	0	2
movies	0	0	0	1	0	1
music[1]	0	2	1	1	1	5
news[2]	0	1	5	2	1	9
own website	1	0	0	0	0	1
reading	1	0	0	0	1	2
school website	0	1	1	1	0	3
search engine	0	0	2	0	0	2
shopping	0	0	2	0	0	2
sports	1	0	1	1	0	3
stars	0	0	1	0	0	1
streaming	1	2	8	2	1	14
streaming games[5]	0	1	2	0	0	3
studying	2	1	0	2	1	6
thematic portal	1	0	1	0	0	2
travelling	1	0	0	0	0	1
video chatting	0	0	0	0	1	1
video platform	0	0	0	0	1	1

[1] These categories were combined to form the category 'music' as reported in section 6.1.3

[2] These categories were combined to form the category 'news' as reported in section 6.1.3

[3] These categories were combined to form the category 'fun/satire' as reported in section 6.1.3

[4] These categories were combined to form the category 'dictionary/translator' as reported in section 6.1.3

[5] These categories were combined to form the category 'gaming' as reported in section 6.1.3

Table B.5: Analysis of qualitative EEQ data on where participants encounter English most frequently

Code	Frequency according to response number			Total	Classification according to major groups[1]	
	1i response 1	1i response 2	1i response 3			
school	39	30	22	91	school	91
at home	0	1	0	1		
abroad/holiday	1	0	2	3		
signs (and public text)	0	2	5	7	linguistic landscape	31
advertisements	2	4	4	10		
auditory signals	0	1	0	1		
public places	0	5	8	13		
sports	0	1	1	2		
internet	48	18	13	79	internet & social media	114
internet/social media	18	10	7	35		
audiovisual media	4	11	4	19	films, series, video clips and other audio visual media	114
online audiovisual media	4	2	0	6		
YouTube/video clips	21	16	8	45		
TV	0	4	3	7		
series	9	5	1	15		
online series	3	1	1	5		
films	2	6	7	15		
cinema	0	1	1	2		
games	5	4	8	17	games & apps	24
apps/programs	3	1	3	7		
phone	5	0	0	5		
chat/messenger	0	2	4	6		
music	13	21	10	44	music	44
radio	0	0	1	1		
reading	1	2	1	4	reading	33
books	1	7	12	20		
news(papers)/magazines	1	2	2	5		
online reading	0	2	2	4		
people	3	7	8	18	other people (incl. chatting)	47
talking to people	0	3	4	7		
tourists	1	4	8	13		
talking to strangers	0	2	1	3		
English in other languages	0	3	3	6		

[1] Empty fields indicate that these codes were not included in any larger category, except for chatting which is including in 'other people'

Table B.6: Frequency of EE activities including summary variables

EE activity[1]	almost never	almost never %	a few times a year	a few times a year %	a few times a month	a few times a month %	a few times a week	a few times a week %	almost daily	almost daily %	N	Sum %[2]
listening to music (overall)	2	1.06	1	0.53	5	2.65	21	11.11	160	84.66	189	100.01
listening to music on a phone/mp3-player	11	5.82	0	0	7	3.7	28	14.81	142	75.13	188	99.46
listening to a Cd or the radio	62	32.8	23	12.17	32	16.93	39	20.63	31	16.4	187	98.93
listening to music on Sportify or other streaming services	39	20.63	5	2.65	19	10.05	36	19.05	89	47.09	188	99.47
watching music videos on the internet	4	2.12	8	4.23	29	15.34	70	37.04	78	41.27	189	100
listening to music at concerts	105	55.56	64	33.86	13	6.88	6	3.17	0	0	188	99.47
watching films (overall)	15	7.94	17	8.99	45	23.81	56	29.63	55	29.1	188	99.47
watching films on TV	87	46.03	38	20.11	40	21.16	17	8.99	5	2.65	187	98.94
watching films on DVD/Blu-ray (overall)	74	39.15	49	25.93	36	19.05	22	11.64	7	3.7	188	99.47
watching films on DVD/Blu-ray with subtitles	105	55.56	47	24.87	21	11.11	9	4.76	5	2.65	187	98.95
watching films on DVD/Blu-ray without subtitles	93	49.21	39	20.63	34	17.99	18	9.52	3	1.59	187	98.94
watching films on the internet (overall)	26	13.76	13	6.88	41	21.69	53	28.04	53	28.04	186	98.41
watching films on the internet (e.g. Netflix, …) with subtitles	90	47.62	16	8.47	34	17.99	28	14.81	17	8.99	185	97.88
watching films on the internet (e.g. Netflix, …) without subtitles	31	16.4	20	10.58	39	20.63	45	23.81	48	25.4	183	96.82
watching films in the cinema	116	61.38	57	30.16	10	5.29	5	2.65	0	0	188	99.48
watching series (overall)	26	13.76	18	9.52	35	18.52	52	27.51	58	30.69	189	100
watching series on TV	122	64.55	26	13.76	28	14.81	9	4.76	3	1.59	188	99.47
watching series on DVD/Blu-ray (overall)	107	56.61	32	16.93	33	17.46	13	6.88	3	1.59	188	99.47
watching series on DVD/Blu-ray with subtitles	135	71.43	30	15.87	17	8.99	4	2.12	1	0.53	187	98.94
watching series on DVD/Blu-ray without subtitles	112	59.26	27	14.29	30	15.87	12	6.35	3	1.59	184	97.36
watching series on the internet (overall)	31	16.4	19	10.05	31	16.4	49	25.93	58	30.69	188	99.47
watching series on the internet (e.g. Netflix, …) with subtitles	84	44.44	23	12.17	36	19.05	27	14.29	18	9.52	188	99.47
watching series on the internet (e.g. Netflix, …) without subtitles	41	21.69	21	11.11	29	15.34	44	23.28	51	26.98	186	98.4
watching programmes on TV	116	61.38	36	19.05	27	14.29	7	3.7	2	1.06	188	99.48
listening to programmes on the radio	151	79.89	16	8.47	14	7.41	3	1.59	2	1.06	186	98.42
watching programmes on the internet	42	22.22	38	20.11	47	24.87	33	17.46	28	14.81	188	99.47

EE activity[1]	almost never	almost never %	a few times a year	a few times a year %	a few times a month	a few times a month %	a few times a week	a few times a week %	almost daily	almost daily %	N	*Sum %[2]*
watching video clips on the internet	2	1.06	2	1.06	21	11.11	52	27.51	109	57.67	186	98.41
listening to audiobooks	171	90.48	9	4.76	4	2.12	0	0	1	0.53	185	97.89
watching plays in a theatre	148	78.31	32	16.93	9	4.76	0	0	0	0	189	100
reading lyrics	19	10.05	22	11.64	54	28.57	51	26.98	43	22.75	189	99.99
translating lyrics or reading translations	57	30.16	44	23.28	41	21.69	24	12.7	22	11.64	188	99.47
reading books (overall)	54	28.57	59	31.22	42	22.22	21	11.11	11	5.82	187	98.94
reading books on paper	60	31.75	59	31.22	43	22.75	16	8.47	6	3.17	184	97.36
reading ebooks	129	68.25	27	14.29	15	7.94	9	4.76	5	2.65	185	97.89
reading articles (overall)	35	18.52	20	10.58	56	29.63	49	25.93	29	15.34	189	100
reading articles on paper	129	68.25	29	15.34	25	13.23	3	1.59	2	1.06	188	99.47
reading articles online	32	16.93	21	11.11	56	29.63	49	25.93	28	14.81	186	98.41
reading information texts (overall)	26	13.76	34	17.99	55	29.1	45	23.81	29	15.34	189	100
reading information texts on paper	114	60.32	42	22.22	22	11.64	9	4.76	1	0.53	188	99.47
reading information texts online (including Wikipedia etc)	27	14.29	35	18.52	53	28.04	44	23.28	29	15.34	188	99.47
reading stories	66	34.92	35	18.52	42	22.22	31	16.4	14	7.41	188	99.47
reading comics	132	69.84	18	9.52	21	11.11	13	6.88	4	2.12	188	99.47
reading blogs/forum entries	63	33.33	36	19.05	46	24.34	24	12.7	20	10.58	189	100
reading e-mails	103	54.5	32	16.93	27	14.29	19	10.05	7	3.7	188	99.47
reading on social media (overall)	11	5.82	5	2.65	16	8.47	40	21.16	114	60.32	186	98.42
reading messages on social media	18	9.52	9	4.76	20	10.58	38	20.11	98	51.85	183	96.82
reading status updates/comments (also about photos) on social media	14	7.41	7	3.7	22	11.64	34	17.99	107	56.61	184	97.35
reading sms/WhatsApp messages	42	22.22	25	13.23	37	19.58	44	23.28	39	20.63	187	98.94
reading subtitles (overall)	59	31.22	29	15.34	40	21.16	36	19.05	25	13.23	189	100
reading offline (overall)	46	24.34	56	29.63	57	30.16	23	12.17	7	3.7	189	100
reading online (overall)	13	6.88	18	9.52	58	30.69	55	29.1	45	23.81	189	100
writing lyrics	132	69.84	18	9.52	13	6.88	14	7.41	11	5.82	188	99.47
writing stories	157	83.07	17	8.99	6	3.17	5	2.65	3	1.59	188	99.47
writing comics	181	95.77	4	2.12	1	0.53	1	0.53	1	0.53	188	99.48
writing blogs/forum entries	168	88.89	10	5.29	8	4.23	1	0.53	2	1.06	189	100
writing e-mails	126	66.67	33	17.46	21	11.11	8	4.23	0	0	188	99.47
writing on social media (overall)	45	23.81	13	6.88	37	19.58	33	17.46	60	31.75	188	99.48
writing messages on social media	57	30.16	14	7.41	40	21.16	26	13.76	48	25.4	185	97.89

EE activity[1]	almost never	almost never %	a few times a year	a few times a year %	a few times a month	a few times a month %	a few times a week	a few times a week %	almost daily	almost daily %	N	Sum %[2]
writing status updates/comments (also about photos) on social media	60	31.75	19	10.05	31	16.4	25	13.23	52	27.51	187	98.94
writing sms/WhatsApp messages	46	24.34	24	12.7	36	19.05	44	23.28	38	20.11	188	99.48
writing lists or notes	88	46.56	41	21.69	22	11.64	16	8.47	11	5.82	178	94.18
writing a diary	162	85.71	8	4.23	4	2.12	2	1.06	1	0.53	177	93.65
speaking English face to face	56	29.63	65	34.39	35	18.52	20	10.58	13	6.88	189	100
speaking English on the phone	118	62.43	38	20.11	19	10.05	11	5.82	3	1.59	189	100
speaking English via Skype or similar internet services	119	62.96	28	14.81	19	10.05	12	6.35	11	5.82	189	99.99
talking to yourself or thinking in English	61	32.28	30	15.87	35	18.52	32	16.93	28	14.81	186	98.41
using English words in other languages	17	8.99	18	9.52	25	13.23	37	19.58	89	47.09	186	98.41
acting in English	182	96.3	5	2.65	2	1.06	0	0	0	0	189	100.01
making video clips yourself	167	88.36	11	5.82	3	1.59	2	1.06	2	1.06	185	97.89
singing (overall)	26	13.76	11	5.82	14	7.41	37	19.58	101	53.44	189	100.01
singing yourself (also karaoke)	82	43.39	19	10.05	28	14.81	29	15.34	31	16.4	189	99.99
singing along to music	25	13.23	11	5.82	14	7.41	39	20.63	99	52.38	188	99.47
playing games (overall)	46	24.34	16	8.47	40	21.16	43	22.75	43	22.75	188	99.47
playing games on a computer or console (overall)	79	41.8	23	12.17	31	16.4	26	13.76	29	15.34	188	99.47
playing games on a phone, tablet or iPod	61	32.28	27	14.29	39	20.63	35	18.52	26	13.76	188	99.48
playing games on a computer or console	82	43.39	22	11.64	32	16.93	26	13.76	26	13.76	188	99.48
playing games with others via internet (Multiplayer Online Games)	109	57.67	11	5.82	23	12.17	19	10.05	25	13.23	187	98.94
chatting in in-game chats while playing a game	18	9.52	6	3.17	12	6.35	13	6.88	30	15.87	79	41.79
speaking to other via VOIP services while playing a game	40	21.16	6	3.17	11	5.82	5	2.65	17	8.99	79	41.79
chatting in English	48	25.4	15	7.94	30	15.87	19	10.05	30	15.87	142	75.13
using English-language apps on your phone	42	22.22	13	6.88	23	12.17	38	20.11	65	34.39	181	95.77
using search engines (e.g. Google) in English	30	15.87	17	8.99	38	20.11	41	21.69	60	31.75	186	98.41

1 The activities are grouped according to the skills involved (see section 6.3) and listed according to the order in the EEQ within groups

2 The percentages given correspond to the percentage of the total number of students in the sample (N=189) to ensure consistency with the figures displayed in section 6.3. Please note that some activities, e.g. chatting in in-game chats while playing a game, have a very high non response rate as this item, for instance, was only answered by participants who play English-language games. In order to highlight the fact that these percentage figures include missing answers, the last column 'Sum %' shows the total percentage accounted for by the five response options, the missing percentage corresponding to missing answers.

Figure B.1: Scatterplot matrix showing relations between influencing factors and mean time spent with EE graphically (lower half) and numerically through Kendall's tau (upper half, * *p* < .05, ** *p* < .01)

Appendix B
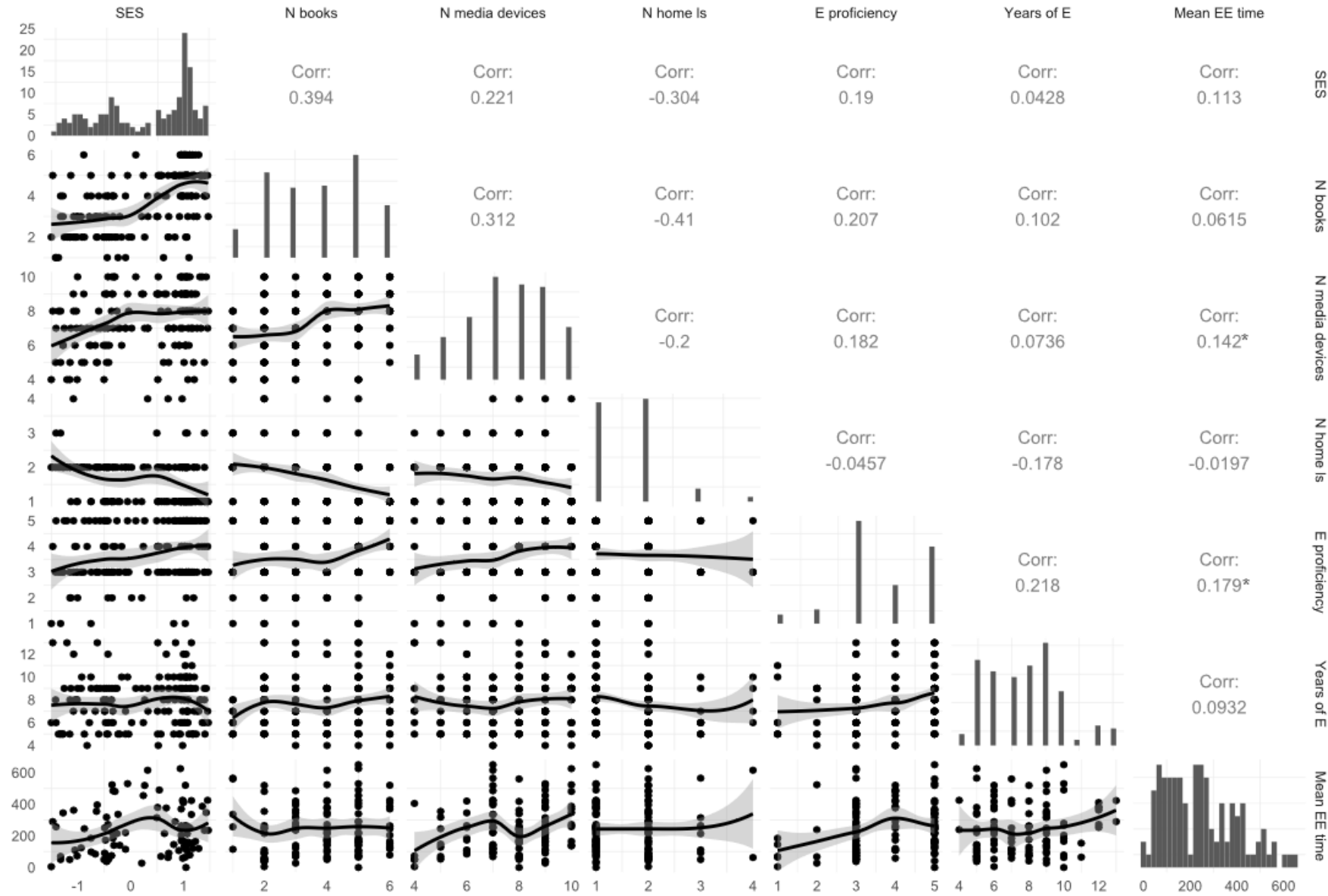
Table B.7: Correlations (Kendall's tau τ) between mean time spent with EE and influencing factors

	EE mean time		
	T	95% CI	p
SES	.11	[-.02, .24]	.110
Number of books available at home	.06	[-.08, .20]	.376
Number of media devices available at home	.14	[.0, .28]	.040
Number of home languages	-.02	[-.17, .13]	.792
Self-assessed English proficiency	.18	[.04, .31]	.013
Length of instruction	.09	[-.05, .23]	.175

Table B.8: Gender differences for individual EE activities

	N	*Mdn female*	95% CI [1]	*Mdn male*	95% CI	*W*	*p*	*r*	95% CI
listening to music on a phone/mp3-player	187	5		5		4283.5	.951	.00	[-.15, .14]
listening to a Cd or the radio	186	3	[3, 4]	2	[1, 3]	5152.5	.008**	-.20	[-.33, -.05]
listening to music on Sportify or other streaming services	187	5	[4, 5]	4	[4, 5]	4740	.153	-.10	[-25., .04]
watching music videos on the internet	188	4	[4, 4]	4	[4, 5]	4361.5	.872	-.01	[-.16, .13]
singing along	187	5		4	[3, 4]	5942	<.001**	-.37	[-.49, -.24]
singing yourself (also karaoke)	188	3	[3, 4]	1	[1, 1]	5874	<.001**	-.33	[-.45, -.19]
listening to music at concerts	187	2	[1, 2]	1		5183.5	.004**	-.21	[-.34, -.07]
reading lyrics	188	4	[3, 4]	3	[3, 4]	5016	.047*	-.14	[-.28, .00]
translating lyrics or reading translations	187	3	[2, 3]	2	[1, 2]	5200	.009**	-.19	[-.33, -.05]
writing lyrics	187	1		1	[1, 1]	4116	.611	-.04	[-.18, .11]
watching films on TV	186	2	[1, 2]	2	[1, 2]	4358.5	.700	-.03	[-.17, .12]
watching films on DVD / Blu-ray **with** subtitles	186	2	[1, 2]	1	[1, 1]	5213	.003**	-.22	[-.36, -.08]
watching films on DVD / Blu-ray **without** subtitles	186	2	[2, 2]	1	[1, 1]	5155	.005**	-.21	[-.34, -.06]

	N	Mdn	[IQR]	Mdn	[IQR]	U	p	r	CI
watching films on the internet **with** subtitles	184	2	[1, 3]	1	[1, 2]	4768.5	.057	-.14	[-.28, .01]
watching films on the internet **without** subtitles	182	3	[3, 4]	4	[3, 4]	3091.5	.007**	-.20	[-.34, -.06]
watching films in the cinema	187	1		1	[1, 2]	3905.5	.251	-.08	[-.23, .06]
watching series on TV	187	1	[1, 1]	1	[1, 1]	4651	.198	-.09	[-.24, .05]
watching series on DVD / Blu-ray **with** subtitles	186	1	[1, 1]	1		4716.5	.078	-.13	[-.27, .02]
watching series on DVD / Blu-ray **without** subtitles	183	1	[1, 2]	1		4567	.118	-.12	[-.26, .03]
watching series on the internet **with** subtitles	187	2.5	[2, 3]	1	[1, 2]	4894	.070	-.13	[-.27, .01]
watching series on the internet **without** subtitles	185	3	[3, 4]	4	[3, 4]	3593.5	.091	-.12	[-.26, .02]
watching programmes on TV	187	1	[1, 1]	1	[1, 2]	3780.5	.137	-.11	[-.25, .04]
listening to programmes on the radio	185	1		1		4216.5	.813	-.02	[-.16, .13]
watching programmes on the internet	187	3	[2, 3]	3	[2, 3.5]	3604	.070	-.13	[-.27, .01]
watching video clips on the internet	185	4	[4, 5]	5		3117.5	.001**	-.24	[-.37, -.01]
making video clips yourself	184	1		1		3865.5	.194	-.10	[-.24, .05]
playing games on a phone, tablet or iPod	187	2	[1, 3]	3	[3, 4]	3001.5	<.001**	-.26	[-.39, -.12]
playing games on a computer or console	187	1		4	[3, 4]	1190.5	<.001**	-.64	[-.72, -.55]
playing games with others via internet (multiplayer online games)	186	1		4	[3, 4]	941	<.001**	-.74	[-.80, -.67]
writing English in in-game chats while playing games	78	1.5	[1, 3]	4	[4, 5]	191.5	.001**	-.39	[-.57, -.18]
speaking English via VOIP services while playing games	78	1		3	[1, 3]	182	<.001**	-.42	[-.59, -.22]
reading books on paper	183	2	[2, 3]	2	[1, 2]	5020	.006**	-.20	[-.34, -.06]
reading ebooks	184	1	[1, 1]	1		4673	.049*	-.15	[-.28, .00]
listening to audiobooks	184	1		1		4271.5	.307	-.08	[-.22, .07]
reading articles on paper	187	1		1	[1, 1]	4309	.887	-.01	[-.15, .13]
reading articles on the internet	185	3	[3, 3]	3	[3, 4]	4008	.610	-.04	[-.18, .11]
reading information texts on paper	187	1	[1, 2]	1	[1, 1]	4912.5	.039*	-.15	[-.29, -.01]
reading information texts on the internet	187	3	[3, 3]	3	[3, 4]	4104	.651	-.03	[-.18, .11]
reading stories	187	3	[2, 3]	2	[1, 2]	5290	.003**	-.22	[-.35, .07]
writing stories	187	1		1		4803	.020*	-.17	[-.31, -.03]
reading comics	187	1		1	[1, 2]	3483.5	.009**	-.19	[-.33, -.05]
writing comics	187	1		1		4432	.133	-.11	[-.25, .03]
reading blogs /forum entries	188	2	[2, 3]	2	[2, 3]	4188	.743	-.02	[-.17, .12]
writing blogs /forum entries	188	1		1		3924.5	.059	-.14	[-.28, .01]

using search engines	185	4	[3, 4]	4	[3, 4]	3851.5	.338	-.07	[-.21, .08]
using apps on your phone	180	4	[3, 4]	4	[4, 4]	3781.5	.535	-.05	[-.19, .10]
reading e-mails	187	1	[1, 1]	2	[1, 2]	3734.5	.119	-.11	[-.25, .03]
writing e-mails	187	1		1	[1, 1]	4058.5	.528	-.05	[-.19, .10]
reading messages on social networks	182	5	[4, 5]	5	[4, 5]	4099.5	.747	-.02	[-.17, .12]
writing messages on social networks	184	3	[3, 4]	3	[2, 3]	4586.5	.148	-.11	[-.25, .04]
reading status updates / comments in English on social networks	183	5	[5, 5]	4	[4, 5]	4921	.006**	-.20	[-.34, -.06]
writing status updates / comments in English on social networks	186	3	[3, 4]	2	[1, 3]	5090	.011*	-.19	[-.32, -.04]
reading sms/messages	186	3	[3, 4]	3	[2, 3]	4715.5	.156	-.10	[-.25, .04]
writing sms/messages	187	3	[3, 4]	3	[2, 4]	4541	.417	-.06	[-.20, .09]
chatting	141	3	[2, 3]	3	[2, 3]	2314	.590	-.05	[-.21, .12]
writing lists or notes	177	2	[1, 2]	1	[1, 2]	4317.5	.084	-.13	[-.27, .02]
writing a diary in English	176	1		1		4260	.001**	-.25	[-.38, -.10]
speaking English via Skype or similar internet services	188	1		1	[1, 2]	3331	.002**	-.22	[-.36, -.08]
speaking English on the phone	188	1	[1, 1]	1	[1, 1]	4220	.789	-.02	[-.16, .12]
speaking English face-to-face	188	2	[2, 2]	2	[2, 2]	4225.5	.822	-.02	[-.16, .13]
talk to yourself or think	185	2	[2, 3]	3	[2, 3]	3723.5	.214	-.09	[-.23, .05]
using English words in other languages	185	4	[4, 5]	5	[4, 5]	4057.5	.767	-.02	[-.17, .12]
watching plays in a theatre	188	1		1		4500.5	.463	-.05	[-.20, .09]
acting yourself in English	188	1		1		4206.5	.415	-.06	[-.20, .09]

p < .05, ** *p* < .01

[1] The empty cells in the columns presenting the 95% confidence intervals for the median in the female and male group could not be calculated due to computational issues: an error that all values in the bootstrapping sample corresponded to the corresponding median and that therefore CIs could not be computed was reported.

Table B.9: Correlations between individual EE activities and overall self-assessed language proficiency and SES

	Self-assessed English proficiency			SES		
	τ	95% CI	p	τ	95% CI	p
listening to music on a phone/mp3-player	.07	[-.07, .21]	.311	.09	[-.05, .23]	.150
listening to a Cd or the radio	.06	[-.07, .18]	.334	.07	[-.04, .20]	.219
listening to music on Sportify or other streaming services	.09	[-.03, .21]	.142	.15	[.03, .27]	.015*
watching music videos on the internet	.06	[-.07, .19]	.366	-.01	[-.14, .10]	.862
singing along	.17	[.05, .29]	.006**	-.02	[-.15, .10]	.702
singing yourself (also karaoke)	.18	[.06, .30]	.004**	-.04	[-.16, .08]	.548
listening to music at concerts	.16	[.03, .28]	.017*	.22	[.10, .32]	.001**
reading lyrics	.11	[-.02, .23]	.070	-.08	[-.19, .03]	.190
translating lyrics or reading translations	.05	[-.08, .17]	.458	-.06	[-.17, .06]	.341
writing lyrics	.12	[-.01, .24]	.066	-.01	[-.12, .11]	.929
watching films on TV	.20	[.08, .31]	.002**	-.04	[-.16, .09]	.528
watching films on DVD / Blu-ray **with** subtitles	.06	[-.08, .19]	.360	.05	[-.07, .16]	.443
watching films on DVD / Blu-ray **without** subtitles	.22	[.11, .34]	< .001**	.21	[.10, .32]	.001**
watching films on the internet **with** subtitles	-.04	[-.16, .08]	.565	-.07	[-.19, .05]	.245
watching films on the internet **without** subtitles	.17	[.04, .29]	.007**	.19	[.07, .30]	.002**
watching films in the cinema	.14	[.00, .26]	.036*	.18	[.05, .30]	.005**
watching series on TV	.08	[-.01, .20]	.223	-.04	[-.17, .09]	.506
watching series on DVD / Blu-ray **with** subtitles	.07	[-.06, .19]	.296	.10	[-.03, .22]	.113
watching series on DVD / Blu-ray **without** subtitles	.15	[.03, .28]	.018*	.21	[.09, .32]	.001**
watching series on the internet **with** subtitles	.01	[-.10, .13]	.832	-.04	[-.16, .08]	.508
watching series on the internet **without** subtitles	.20	[.09, .32]	.001**	.17	[.06, .28]	.005**
watching programmes on TV	.12	[-.01, .25]	.069	-.07	[-.20, .05]	.237
listening to programmes on the radio	.14	[.02, .27]	.030*	.13	[.01, .23]	.053
watching programmes on the internet	.22	[.10, .33]	< .001**	.04	[-.07, .17]	.479
watching video clips on the internet	.17	[.04, .29]	.011*	.03	[-.10, .16]	.599
making video clips yourself	.12	[-.02, .25]	.076	.05	[-.06, .16]	.418
playing games on a phone, tablet or iPod	-.01	[-.13, .12]	.902	.05	[-.07, .16]	.435
playing games on a computer or console	.10	[-.03, .22]	.113	.10	[-.03, .21]	.111
playing games with others via internet (multiplayer online games)	.06	[-.06, .19]	.327	.15	[.02, .27]	.018*

Appendix B

writing English in in-game chats while playing games	.13	[-.07, .33]	.179	.00	[-.20, .19]	1
speaking English via VOIP services while playing games	.23	[.02, .42]	.018*	.10	[-.11, .31]	.292
reading books on paper	.33	[.22, .44]	< .001**	.13	[.01, .25]	.031*
reading ebooks	.17	[.04, .29]	.010*	.16	[.06, .27]	.010*
listening to audiobooks	.08	[-.06, .21]	.240	.19	[.10, .28]	.003**
reading articles on paper	.22	[.11, .34]	.001**	.17	[.06, .28]	.007**
reading articles on the internet	.24	[.12, .36]	< .001**	.04	[-.08, .15]	.479
reading information texts on paper	.22	[.10, .35]	.001**	.14	[.01, .26]	.030*
reading information texts on the internet	.32	[.20, .43]	< .001**	.10	[-.02, .21]	.085
reading stories	.22	[.10, .34]	< .001**	-.01	[-.12, .11]	.896
writing stories	.10	[-.05, .23]	.139	.12	[.02, .23]	.056
reading comics	.09	[-.03, .22]	.159	.02	[-.10, .15]	.697
writing comics	-.01	[-.12, .12]	.880	.07	[-.01, .15]	.311
reading blogs /forum entries	.16	[.04, .28]	.010*	.07	[-.04, .18]	.245
writing blogs /forum entries	.08	[-.06, .21]	.230	.05	[-.08, .17]	.443
using search engines	.32	[.21, .43]	< .001**	.23	[.11, .34]	< .001**
using apps on your phone	.16	[.02, .27]	.014*	.11	[.00, .23]	.062
reading e-mails	.26	[.15, .38]	< .001**	.13	[.01, .25]	.031*
writing e-mails	.29	[.18, .40]	< .001**	.16	[.05, .28]	.009**
reading messages on social networks	.22	[.10, .34]	.001**	.03	[-.10, .15]	.635
writing messages on social networks	.24	[.12, .36]	< .001**	.09	[-.02, .21]	.124
reading status updates / comments in English on social networks	.16	[.03, .28]	.014*	.02	[-.01, .15]	.769
writing status updates / comments in English on social networks	.22	[.10, .33]	< .001**	.09	[-.03, .21]	.143
reading sms/messages	.15	[.04, .27]	.013*	.05	[-.08, .17]	.414
writing sms/messages	.17	[.04, .29]	.006**	.07	[-.05, .19]	.210
chatting	.25	[.12, .38]	.001**	.06	[-.07, .20]	.354
writing lists or notes	.27	[.14, .38]	< .001**	.06	[-.06, .18]	.333
writing a diary in English	.15	[.01, .28]	.027*	.03	[-.10, .15]	.688
speaking English via Skype or similar internet services	.29	[.17, .41]	< .001**	.11	[-.01, .22]	.086
speaking English on the phone	.19	[.06, .30]	.004**	.13	[.00, .25]	.039*
speaking English face-to-face	.25	[.14, .37]	< .001**	.09	[-.02, .19]	.143
talk to yourself or think	.37	[.26, .48]	< .001**	.16	[.04, .26]	.008**
using English words in other languages	.19	[.07, .31]	.002**	.13	[.02, .24]	.028*
watching plays in a theatre	.21	[.10, .32]	.001**	.15	[.04, .26]	.022*
acting yourself in English	-.06	[-.15, .06]	.343	.13	[.03, .22]	.041*

* $p < .05$, ** $p < .01$

Table B.10: Summary statistics for the comparison of the V_YesNo and *h×CJ%* scores to previously proposed correction formulae

Scoring formulae	N	Min	Max	Mdn	95% CI	M	95% CI	SD
h (Number of correct responses)	175	0.180	0.860	0.520	[0.480, 0.550]	0.526	[0.504, 0.548]	0.148
h-f (Hit rate minus false alarm rate)	175	0.160	0.820	0.430	[0.400, 0.450]	0.448	[0.428, 0.469]	0.134
Correction for guessing (*cfg*) (e.g. Meara & Buxton 1987)	175	0.163	0.854	0.480	[0.452, 0.495]	0.488	[0.467, 0.510]	0.145
Δ*m* (Meara 1992)	175	-0.130	0.808	0.331	[0.297, 0.360]	0.347	[0.321, 0.374]	0.186
I_{SDT} (Huigbregste, Admiraal & Meara 2002)	175	0.318	0.872	0.560	[0.542, 0.581]	0.577	[0.561, 0.595]	0.114
Logistic Weighting Function (Meara & Miralpeix 2017)	175	0.160	0.845	0.465	[0.445, 0.490]	0.483	[0.463, 0.505]	0.141
h×CJ%	174	0.144	0.751	0.365	[0.342, 0.396]	0.389	[0.370, 0.409]	0.131

Information on the six previously published scoring formulae for Yes/No tests has been summarized in Table 5.6 in section 5.3.3.3. To allow for easier comparison, the proportion of hits and false alarms was used in calculations and the scores based on the S-shaped logistic weighting function by Meara and Miralpeix (2017) formula, which yields an estimate of the number of words known among the 10,000 most frequent words, were divided by 10,000 to get a comparable percentage score. To convert the hits adjusted by the proportion of correct judgements into a comparable figure, the scores were also divided by the total number of 10,000 words analogous to the V_YesNo score.

Figure B.2: Line graph comparing seven adjustment methods for Yes/No tests: six published scoring formulae and correction based on the number of correct judgements in relation to a translation task
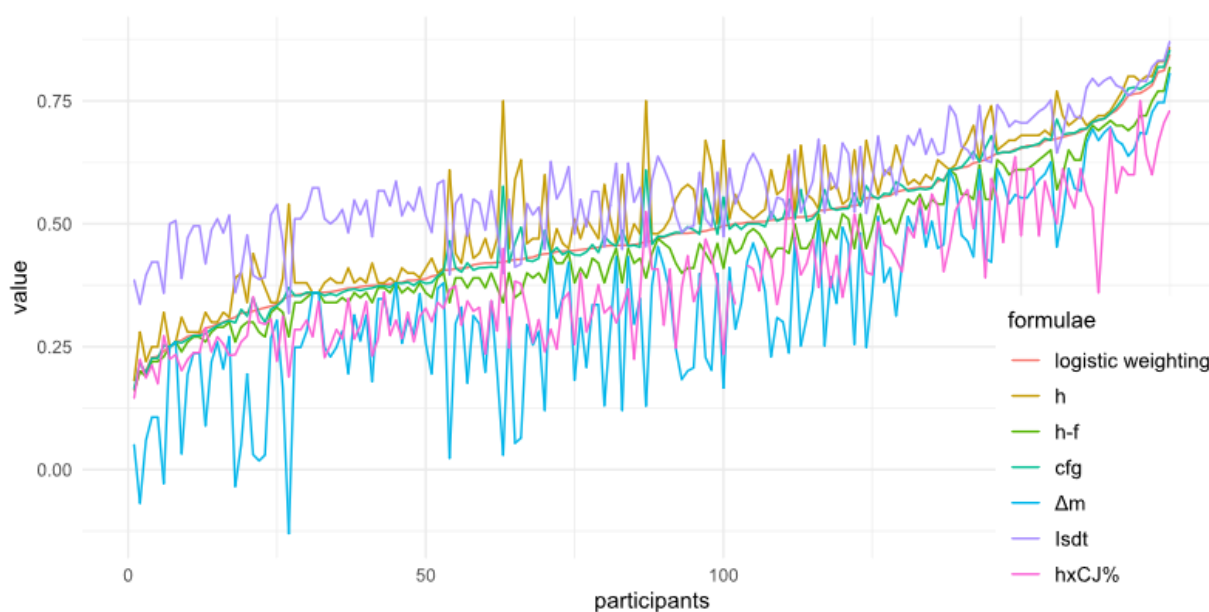
Figure B.3: Boxplot (left) displaying median and interquartile range and dot plot (right) showing mean and standard deviation (red error bar) of *h×CJ%* scores according to gender (*N* = 141)

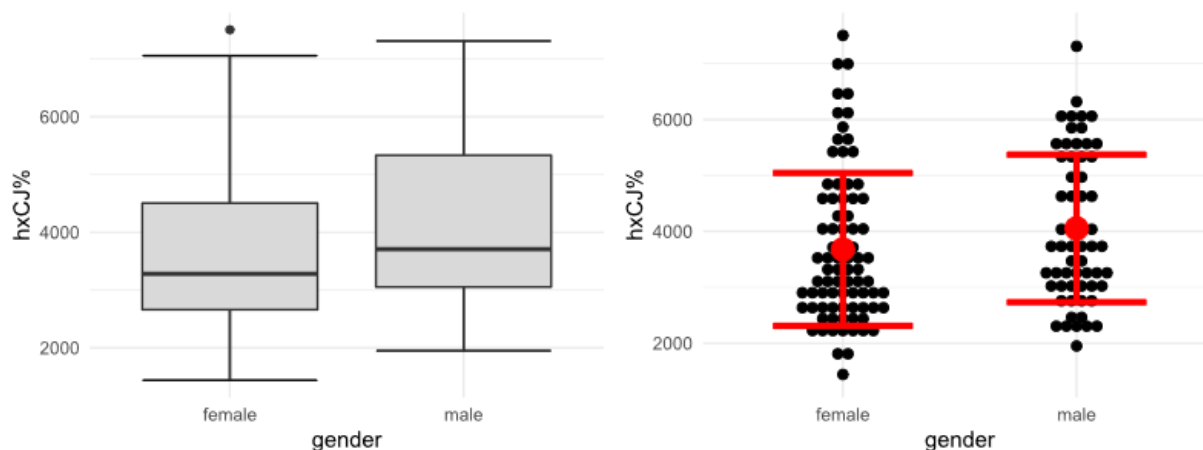Table B.11: Summary statistics for a gender difference in receptive vocabulary scores based on *h×CJ%*

	female participants	male participants
N	81	60
Median	3280	3708
95% CI	[2880, 3610]	[3300, 4328]
Mean	3677.22	4052.67
95% CI	[3404, 3992]	[3720, 4385]
SD	1367.47	1320.69

A two-tailed Wilcoxon rank-sum test (*W* = 1978.5, *p* = .060, *r* = -.16 [-.32,.01]) showed that the difference in receptive vocabulary size based on hit rates adjusted by the proportion of correct judgements between female and male students was not statistically significant.

Table B.12: Correlations (Kendall's tau *τ)* between receptive vocabulary scores based on *h×CJ%* and influencing factors

Influencing factors	*τ*	*95% CI*	*p*
SES (summary variable)	.19	[.07, .31]	.002
Number of books at home	.24	[.14, .34]	< .001
Access to media at home	.21	[.06, .29]	< .001
Number of home languages	-.14	[-.28, .00]	.038
Years spent learning English	.23	[.11, .35]	< .001
EE median score	.20	[.07, .32]	.002

Figure B.4: Scatterplot matrix showing relations between influencing factors and *hxCJ%* score graphically (lower half) and numerically through Kendall's tau (upper half, * $p < .05$, ** $p < .01$)
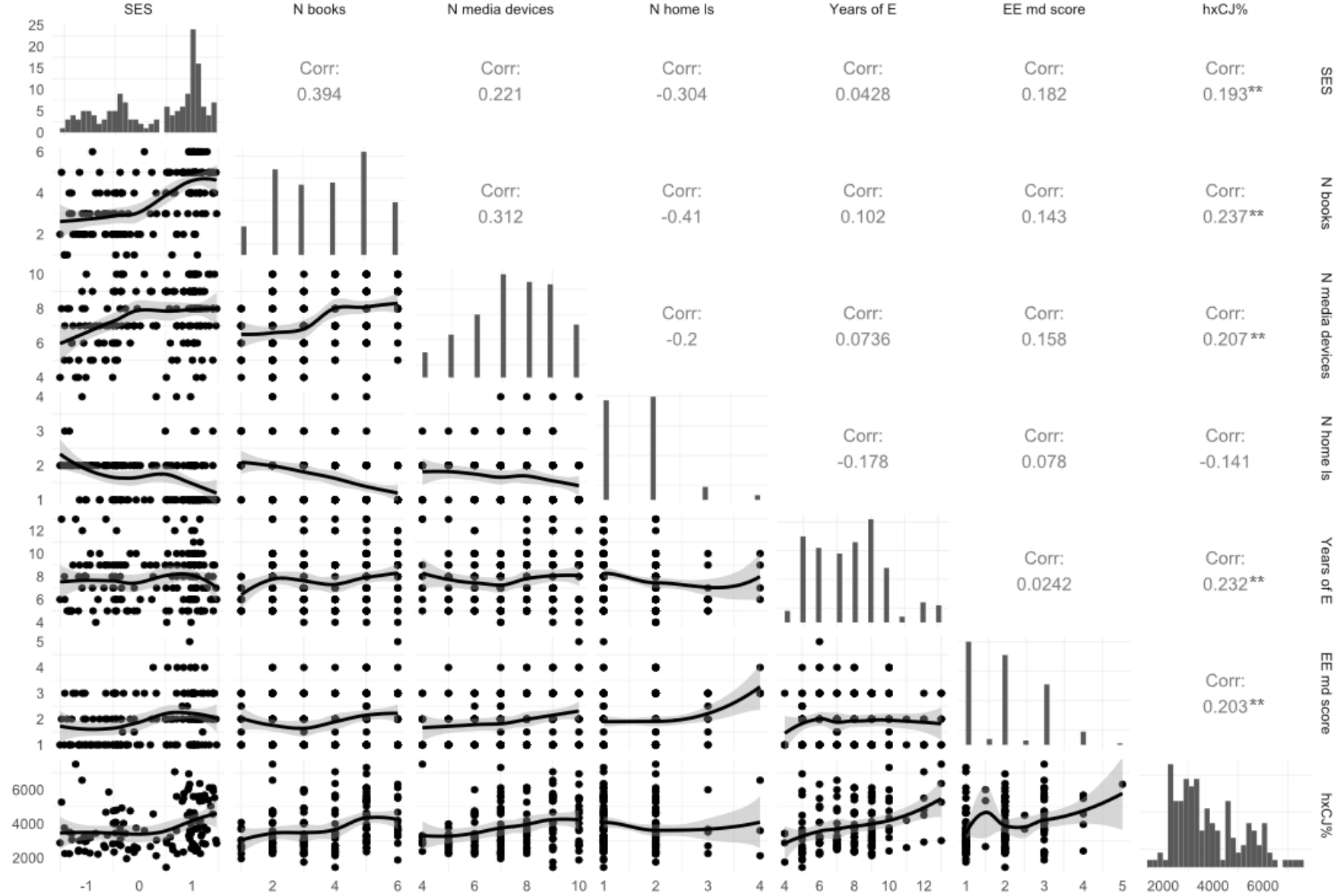
Figure B.5: Boxplot (left) displaying median and interquartile range and dot plot (right) showing mean and standard deviation (red error bar) of *h×CJ%* scores according to EE extreme groups based on EE median score (*N* = 95)
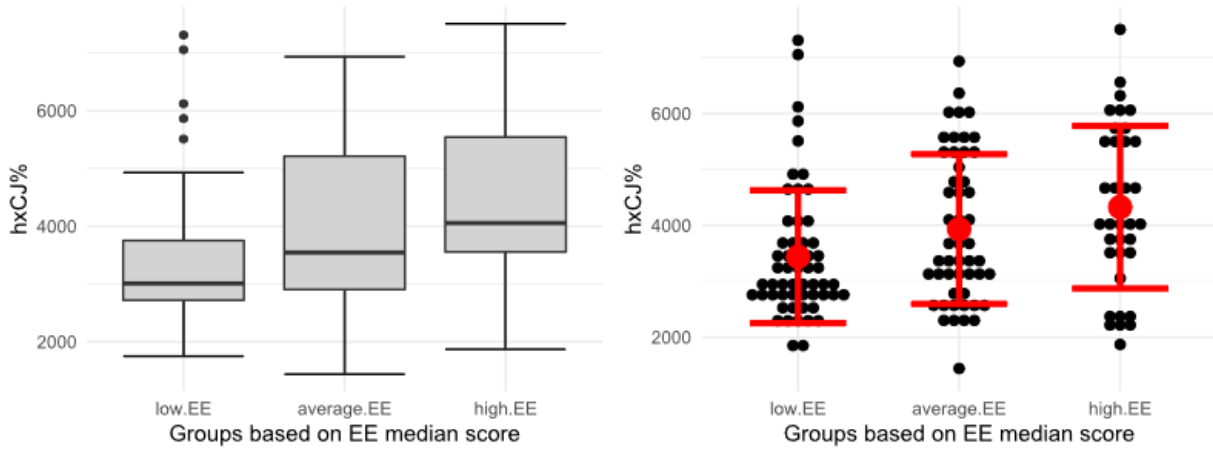

Figure B.6: Boxplot (left) displaying median and interquartile range and dot plot (right) showing mean and standard deviation (red error bar) of *h×CJ%* scores according to EE extreme groups based on EE mean time as measured in the extramural English online diary (*N* = 90)
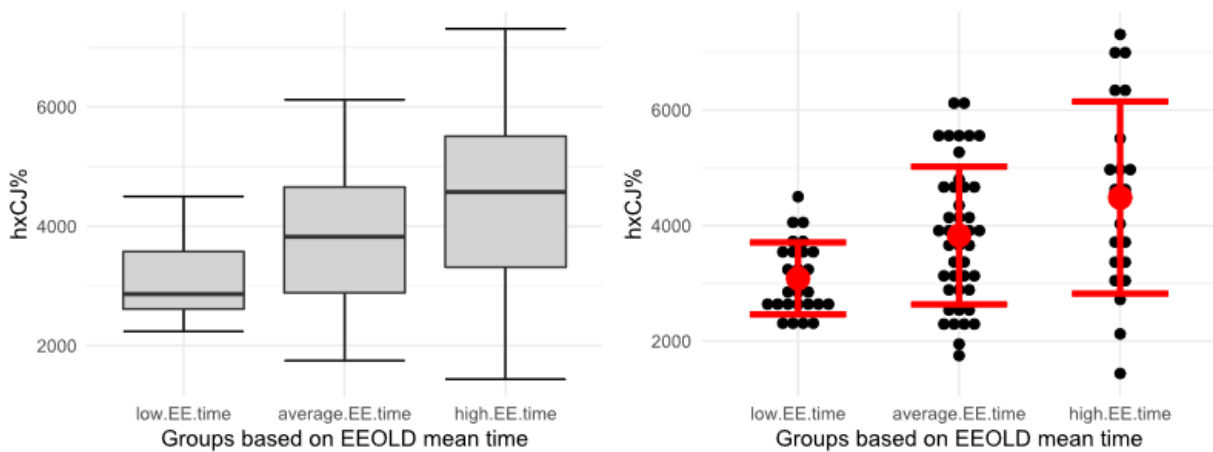
Table B.13: Summary statistics for *h×CJ%* extreme group analysis based on EE median score and EE mean time

	Extreme groups based on EE median score			Extreme groups based on EE mean time		
	low EE group	average EE group	high EE group	low EE group	average EE group	high EE group
N	56	50	36	26	43	21
Median	3010	3545	4055	2865	3825	4575
95% CI	[2850, 3340]	[3200, 4500]	[3715, 4760]	[2665, 3450]	[3120, 4080]	[3315, 5040]
Mean	3438.84	3934.3	4325.42	3085.19	3828.84	4485.24
95% CI	[3161, 3799]	[3584, 4303]	[3862, 4795]	[2869, 3349]	[3481, 4174]	[3798, 5139]
SD	1187.87	1339.04	1453.93	623.02	1192.75	1662.09

A Kruskal-Wallis test (H = 9.91, df = 2, p = .007) showed that the differences in receptive vocabulary size are statistically significant between the three EE groups based on the EE median score. Post-hoc comparisons with one-tailed pairwise Wilcoxon rank-sum tests using the Bonferroni correction indicate that the differences between the low EE group and the high EE group (p = .004) and the low EE group and the average EE group (p = .007) are statistically significant. The difference between the average and the high EE group is not statistically significant (p = .323)

For the groups based on EE mean time, a second Kruskal-Wallis test (H = 12.27, df = 2, p = .002) showed that the differences between groups are statistically significant as well. Again, post-hoc comparisons with one-tailed pairwise Wilcoxon rank-sum tests using the Bonferroni correction show that the difference between the low EE group and the average EE group (p = .015) as well as the high EE group (p = .002) are statistically significant, whereas the difference between the average and the high EE group (p = .208) is not.

Table B.14: Summary statistics for the receptive vocabulary size of the sub-sample engaging in niche activities based on *h×CJ%*

	Sub-sample engaging in niche activities	Remaining participants
N	29	113
Median	3825	3420
95% CI	[2880, 3610]	[3185, 3702]
Mean	4168	3753
95% CI	[3404, 3992]	[3538, 4008]
SD	1557.23	1287.04

A one-tailed Wilcoxon rank-sum test (W = 1904.5, p = .090) indicates that the difference in mean receptive vocabulary size as measured by *h×CJ%* is not statistically significant with a small effect (r = -.14 [-.31, .02]).

Figure B.7: Dot plots comparing the *h×CJ%* scores in a sub-sample of participants engaging in niche EE activities (left) and the remaining participants (right) showing the mean and standard deviation (red error bar)
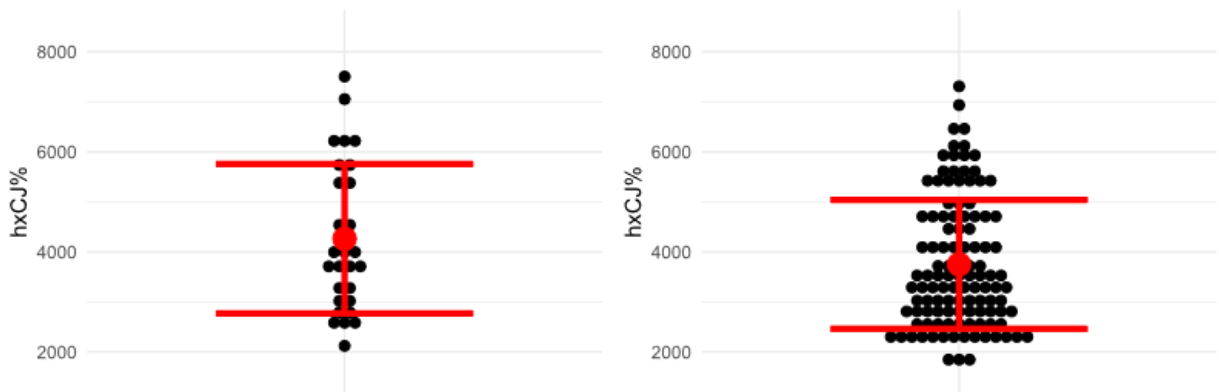
Figure B.8: Scatterplots of the relationship between the predictors included in the regression model and the outcome variable *h×CJ%* (number of hits adjusted by proportion of correct judgements)
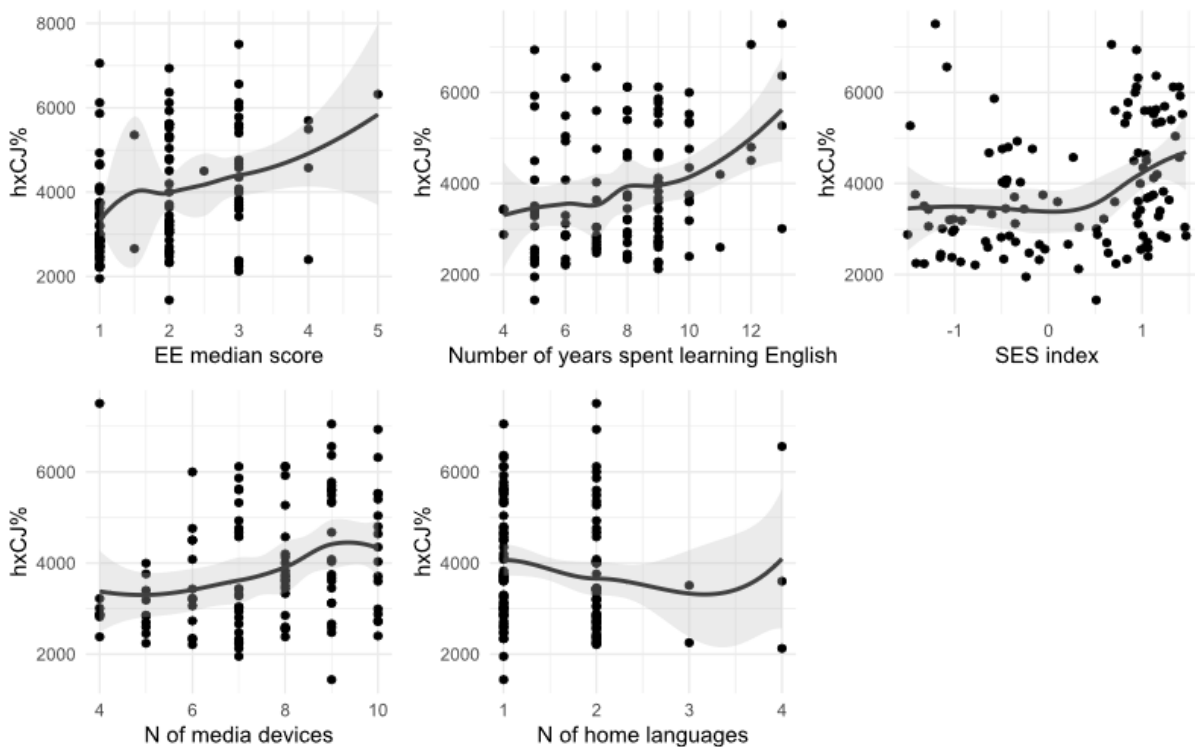
Figure B.9: Diagnostic plots for the linear regression model: residuals vs fitted values (upper left), normal Q-Q plot of residuals (upper right), residuals vs leverage plot (lower left) and Cook's distance (lower left) for model with *h×CJ%*
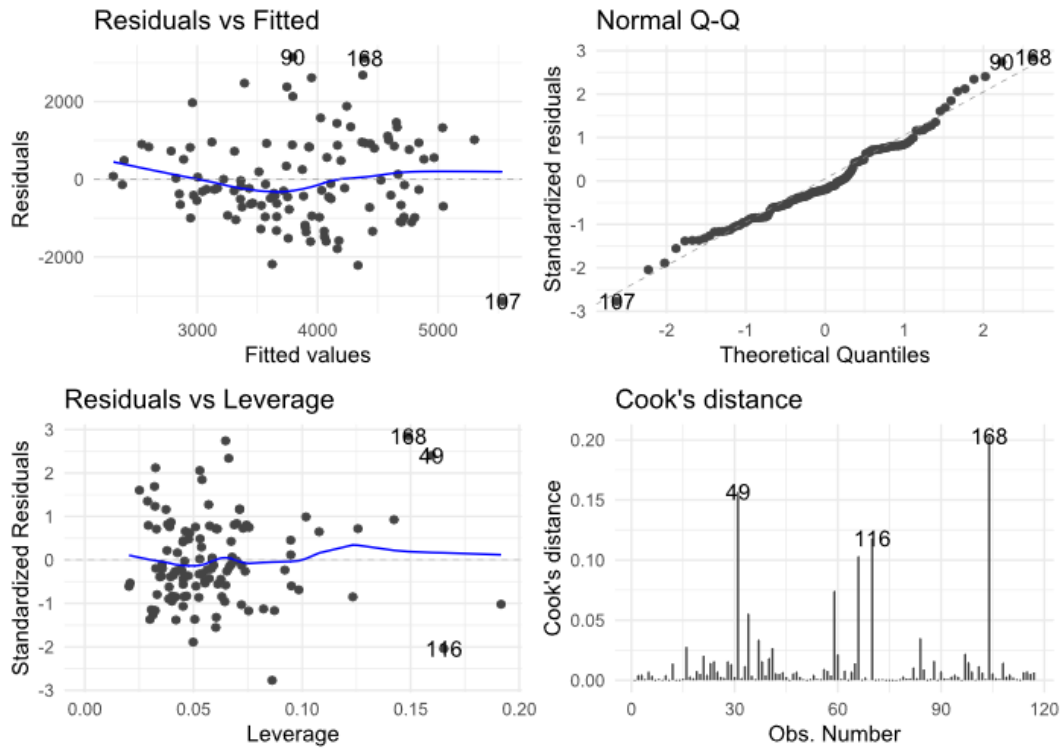
Table B.15: Thematic fields identified in the Lex30 schoolbook analysis

Food [1]

alcohol	cookie	ketchup	peel	steak
avocado	corn	kumquat	pineapple	stew
beet	cornflake	leftover	popcorn	sugar
blueberry	cranberry	lemon	protein	sushi
boil	cucumber	lemonade	raspberry	syrup
burger	digest	mash	raw	vitamin
calorie	filet	mayonnaise	sandwich	watermelon
candy	flavor	meat	soy	wedge
carbohydrate	fry	melon	spaghetti	yummy
cherry	ginger	pan	spatula	
chopsticks	grapefruit	peach	spice	
cocktail	gravy	pear	starfruit	

Medicine

adrenalin	blind	epidemic	medication	surgery
adrenaline	blood	fever	obesity	symptom
amnesia	caries	glasses	pandemic	toxic
amputation	cast	heal	pharmacy	trauma
antibiotic	cavity	hygiene	plague	vaccine
asthma	dentist	immune	pneumonia	virus
autism	disable	infection	recovery	wound
bacteria	depression	injection	rheumatism	

Beauty & clothes

belt	dresscode	laundry	robe	textile
bra	fabric	lotion	scarf	tie
bracelet	garment	panties	scent	ugly
brand	handsome	pants	sneaker	underwear
brush	jeans	pullover	sweater	vintage
clothing	knit	purse	sweatshirt	wellness
cotton	label	pyjama	tanktop	

Crime, terror & war

aggression	confront	loot	scam	terrorist
aggressive	confrontation	mafia	shot	violence
assassination	defeat	martial	shotgun	weapon
assault	defend	massacre	sniper	
bloodthirsty	defense	offence	stalk	
combat	explosive	offensive	terror	
conflict	hostage	prisoner	terrorism	

Science

acid	biological	element	hydrogen	optic
antidote	chart	evolution	hypothesis	paleontology
archeology	chemical	fluid	innovation	quantum
astrobiology	clone	galaxy	laboratory	radioactive
astronomy	compound	geographic	molecule	topographic
atom	discovery	geometry	nuclear	toxic

Feelings

anger	detest	grin	motivation	positive
boredom	enthusiasm	hopeful	negative	ridiculous
comfort	exhaustion	humor	numb	sensation
console	festive	intimate	optimism	sorrow
cosy	frustration	loss	optimistic	stressful
despair	grief	misery	passion	trauma

Character

candid	intelligence	patience	stupidity	violent
creative	inventive	personality	submissive	wisdom
dumb	lazy	selfish	successful	wise
faithful	loyal	smug	supportive	witty
ignorant	loyalty	stamina	thankful	
inquisitive	optimistic	stubborn	vile	

Body

ass	brain	dick	jaw	nerve
bite	breast	digest	joint	pee
blood	breathe	fingernail	lung	reflex
boob	butt	fist	mental	thigh
booger	cock	itch	muscle	

Death

cemetery	deathbed	graveyard	massacre	shot
choke	decease	grief	maul	slay
coffin	drown	immortal	mutilate	starvation
corpse	grave	knell	perish	strangle

Politics

anarchy	corrupt	dictatorship	mob	resolution
authoritarian	county	feminism	parliament	revolution
civilization	coup	hegemony	politician	treaty
conspiracy	dictator	legal	regime	

Economy

advertisement	corporation	finance	payment	property
auction	corrupt	globalization	product	prosperity
broker	economy	import	production	trademark
capitalism	fairtrade	merchant	profit	

Fantasy

alchemy	enchant	invisible	supernatural	witchcraft
antidote	extraterrestrial	monster	unicorn	wizardry
dwarf	fairy	portion	wand	
elf	immortal	sorcery	warlock	

Historical warfare

conquer	foe	saber	slay
defeat	fortress	savage	sword
defend	kingdom	shield	vengeance
defense	pickaxe	siege	warrior

Sports

adrenalin	contest	paragliding	skydiving
adrenaline	jog	polo	soccer
billiards	karate	rugby	tennis
championship	kickbox	skatepark	volleyball

Furniture

bench	couch	lampstand	pillow	stove
bookshelf	cushion	mattress	shelf	wallpaper
closet	decoration	oven	stool	

Pejorative terms

ass	dick	idiotic	pervert	snitch
buster	dumb	knucklehead	pigheaded	
creep	idiot	moron	retard	

Religion & beliefs

advent	faith	preach	sanctuary
destiny	fate	relic	sin
eternal	pray	ritual	worship

Nature

earth	harvest	rainbow	snow
environment	lightning	raindrop	soil
habitat	mud	sand	tsunami

Media

advertisement	entertain	interview	talkshow
blockbuster	flatscreen	sitcom	video
documentary	genre	subtitle	

Computers & technology

browse	chart	internet	laser	server
cable	flatscreen	laptop	network	tablet

Work & jobs

broker	poet	promotion	stewardess
dentist	politician	steward	teamwork

Animals

hedgehog	panda	shark	turtle
moth	racoon	sloth	worm

Drugs

addict	alcohol	dope
addiction	cigarette	weed

[1] As described in section 6.4.9, 27 words were assigned a primary and a secondary thematic field, thus this table includes repeated word types in different thematic fields.

Deutsche Zusammenfassung

Aufgrund des Status als derzeit unangefochtene Weltsprache hat Englisch in den letzten Jahrzehnten in vielen Bereichen des Alltags Einzug gehalten. Dies gilt im Besonderen auch für Länder wie Österreich, in denen Englisch traditionell als Fremdsprache gesehen wird. Aus diesem Grund kommen SprachenlernerInnen heutzutage auch während ihrer Freizeit vielfach mit der Zielsprache in Kontakt. Ein internationales Forschungsfeld, das sich dem Untersuchungsgegenstand des informellen außerschulischen Fremdsprachenlernens widmet, ist derzeit im Entstehen begriffen. In Europa konzentriert sich die bisherige Forschung jedoch stark auf Länder, in denen Fernsehsendungen in Originalsprache mit Untertiteln ausgestrahlt werden; dies ist insofern von Bedeutung, als der ständige Kontakt mit Englisch durch Fernsehprogramme nachweislich einen Einfluss auf den Spracherwerb hat. In Ländern, in denen fremdsprachige Inhalte standardmäßig synchronisiert werden, gibt es hingegen vergleichsweise wenige Studien zum informellen Spracherwerb.

Dieses Dissertationsprojekt hat sich zum Ziel gesetzt, erstmals systematisch die Beschäftigung mit außerschulischem Englisch unter Wiener AHS-SchülerInnen zu beschreiben und einen möglichen Einfluss auf den Spracherwerb am Beispiel des Wortschatzes zu untersuchen. Insgesamt nahmen 201 SchülerInnen der zehnten Schulstufe an der Mixed-Methods-Studie teil. In der quantitativen Studienphase wurden Daten über Häufigkeit und Dauer des Kontakts mit Englisch außerhalb des Unterrichts mit Hilfe eines detaillierten Fragebogens und eines Online-Sprachtagebuchs gesammelt und der Umfang des rezeptiven und produktiven Wortschatzes durch zwei Vokabeltests bestimmt. Zusätzlich wurden in der nachfolgenden qualitativen Studienphase noch Fokusgruppeninterviews mit 30 SchülerInnen in sechs Gruppen durchgeführt, um Einblicke in die Perspektive der Jugendlichen zu bekommen, da sie zweifelsohne die wichtigsten Akteure in informellen Lernprozessen sind.

Die Ergebnisse zeigen, dass die außerschulische Beschäftigung mit Englisch von einigen sehr häufigen Aktivitäten geprägt wird, gleichzeitig aber große Diversität und ein hohes Maß an Individualisierung aufweist. Die überwiegende Mehrheit der teilnehmenden SchülerInnen hat täglich außerschulischen Kontakt mit Englisch, im Durchschnitt sind Jugendliche etwa vier Stunden pro Tag von Englisch umgeben. In Bezug auf die Größe des englischen Vokabulars zeigen die errechneten Regressionsmodelle einen positiven Effekt der außerschulischen Beschäftigung mit Englisch auf den rezeptiven, nicht aber den produktiven Wortschatz. Weitere Analysen deuten jedoch darauf hin, dass auch produktives Vokabelwissen durch englische Freizeitaktivitäten erworben wird. Die Auswertung der Interviewdaten gibt wertvolle Einblicke in die Sichtweisen der SchülerInnen und zeigt, dass die TeilnehmerInnen die außerschulische Beschäftigung mit Englisch als förderlich für den Sprach- und Wortschatzerwerb empfinden, wenngleich sie den schulischen Englischunterricht in der Unterstufe als Basis für ihre englischen Freizeitaktivitäten ansehen. Im Vergleich zu früheren Studien ergeben sich aus den Resultaten dieses Projekts zwei wichtige Schlussfolgerungen: Einerseits ist der Unterschied

zwischen Synchronisationsländern und jenen, die originalsprachliche Fernsehinhalte bevorzugen, in Bezug auf die Häufigkeit und Dauer des Kontakts von Jugendlichen mit außerschulischem Englisch vernachlässigbar. Andererseits scheint der frühe Kontakt mit Englisch durch Fernsehsendungen aber einen wesentlichen Einfluss auf den Verlauf und die Ergebnisse informellen Spracherwerbs zu haben.

Insgesamt leistet die Studie durch den detaillierten Vergleich der verschiedenen Konzeptualisierungen des Forschungsgegenstands, durch die Einführungen methodischer Innovationen im Bereich des Mixed-Methods-Studiendesigns und der Messung des Wortschatzes, sowie durch die empirische Erforschung eines neuen Forschungskontextes aus mehreren unterschiedlichen Perspektiven einen signifikanten Beitrag zum neuen Forschungsgebiet des informellen Sprachenlernens.