



universität
wien

DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

„An analysis of how academic vocabulary development relates to vocabulary learning strategies and the biographical background of English students at the University of Vienna“

verfasst von / submitted by

Katharina Ghamarian, BA

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Magistra der Philosophie (Mag.phil.)

Wien, 2017 / Vienna, 2017

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 190 344 333

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Lehramtsstudium Unterrichtsfach Englisch
Unterrichtsfach Deutsch

Betreut von / Supervisor:

Mag. Dr. Helen Heaney, BA

*"If plan A did not work.
Don't worry the alphabet has 25 more letters."*

Claire Cook

Acknowledgments

Writing a diploma thesis is a challenging and time-consuming task, which requires a lot of inspiration, motivation and perseverance. However, with the support of the right people all challenges are easier to take. Therefore, I would like to seize this opportunity to express my gratitude.

First of all, I would like to thank my supervisor Dr. Helen Heaney, BA who has invested a lot of time and passion into me and my project and has been a great help throughout the development this thesis and has accompanied me from the beginning of my studies until the end.

Secondly, I want to send all my love to my parents who have always supported me and, even more importantly, always believed in me. Without their generosity, time, passion and empathy this thesis and my studies would not have been possible.

The same accounts for my boyfriend Burkart Datler who never got tired of cheering me up and who has accompanied me during my whole studies.

I would also like to express my gratitude to Sandra for her support and the time she invested. The same accounts for my friends who encouraged me whenever they could. Special thanks go to Raffael who did not only read (and print) my thesis several times but demonstrated a lot of passion, interest and input for my work.

Moreover, I would like to thank all lecturers and students who were willing to participant in my study for their time and support. Another word of thanks also goes to Mag. Dr. Christine Simone Sing who always had an open ear for my questions and who has somehow managed to awaken my interest and passion for Linguistics.

Last but not least, I would like to thank all people who have shared my excitement and who have crossed their fingers for me during my studies.

Thank you very much.

Table of contents

| | |
|--|----|
| List of abbreviations | i |
| List of tables | i |
| List of figures..... | ii |
| | |
| 1. Introduction | 1 |
| 2. Literature review | 3 |
| 2.1. Vocabulary in academic writing | 3 |
| 2.1.1. High- and low-frequency words | 4 |
| 2.1.2. Academic vocabulary..... | 5 |
| 2.1.3. Technical vocabulary..... | 6 |
| 2.2. Word lists | 8 |
| 2.2.1. Representativeness..... | 9 |
| 2.2.2. Frequency range and dispersion | 9 |
| 2.2.3. Ways of counting..... | 10 |
| 2.2.4. Collocations | 11 |
| 2.2.5. Range of information..... | 13 |
| 2.2.6. Academic and high-frequency word lists..... | 14 |
| 2.3. Corpus analysis programs..... | 21 |
| 2.4. Vocabulary learning..... | 24 |
| 2.4.1. Aspects of word knowledge | 25 |
| 2.4.2. Learning styles and learning strategies | 29 |
| 2.4.2.1. Language learning strategies | 31 |
| 2.4.2.2. Taxonomy of vocabulary learning strategies..... | 34 |
| 2.4.2.2.1. Planning strategies | 35 |
| 2.4.2.2.2. Recording strategies | 38 |
| 2.4.2.2.3. Discovery strategies | 42 |
| 2.4.2.2.4. Consolidation strategies..... | 46 |
| 2.4.2.2.5 Self-management strategies | 50 |
| 2.5. Biographical background | 53 |
| 2.5.1. Gender..... | 53 |
| 2.5.2. School type | 55 |
| 2.5.3. Bilingualism..... | 58 |
| 2.5.4. Experience abroad..... | 59 |

| | |
|--|----|
| 3. Research | 60 |
| 3.1. Research questions..... | 60 |
| 3.2. Participants | 60 |
| 3.3. Methodology..... | 61 |
| 3.4. Results quantitative study..... | 62 |
| 3.4.1. Vocabulary development between the different points of measurement..... | 62 |
| 3.4.2. Proficiency levels | 68 |
| 3.4.3. Biographical background | 71 |
| 3.4.4. Questionnaire..... | 74 |
| 3.4.5. Vocabulary learning strategies use and the development of vocabulary use..... | 77 |
| 3.4.6. Vocabulary learning strategies and biographical data | 80 |
| 3.5. Results qualitative study | 82 |
| 3.6. Discussion..... | 84 |
| 4. Conclusion | 89 |
| 5. References | 90 |
| 6. Appendix | 96 |
| 6.1. Abstract | 96 |
| 6.2. Zusammenfassung | 97 |

List of abbreviations

| | |
|-----|--|
| AVL | Academic Vocabulary List (Gardner & Davies 2014) |
| AWL | Academic Word List (Coxhead 2000) |
| DET | Determination strategies |
| EFL | English as a foreign language |
| ESP | English for specific purposes |
| GSL | General Service List (West 1953) |
| L1 | First language; mother tongue |
| L2 | Second language |
| L3 | Third language |
| MEM | Memory strategies |
| SOC | Social strategies |

List of tables

| | |
|---|----|
| Table 1: Identifying technical words (Chung & Nation 2003: 105)..... | 7 |
| Table 2: Vocabulary learning strategies..... | 53 |
| Table 3: Mean of vocabulary types used by students..... | 62 |
| Table 4: Mauchly's Sphericity test academic vocabulary..... | 63 |
| Table 5: Significance test for academic vocabulary..... | 64 |
| Table 6: Pairwise comparison of academic vocabulary..... | 64 |
| Table 7: Mauchly's Test of Sphericity high-frequency (1-2000)..... | 65 |
| Table 8: Significance test for high-frequency vocabulary development (1-2000)..... | 65 |
| Table 9: Pairwise comparison of high-frequency vocabulary (1-2000)..... | 66 |
| Table 10: Significance test for high-frequency vocabulary development (2000-4000)... | 66 |
| Table 11: Pairwise comparison of high-frequency vocabulary (2000-4000)..... | 67 |
| Table 12: Sphericity Test for low-frequency and technical vocabulary..... | 67 |
| Table 13: Significance test for low-frequency and technical vocabulary..... | 67 |
| Table 14: Pairwise comparison low-frequency and technical vocabulary..... | 68 |
| Table 15: Cronbach's Alpha..... | 75 |
| Table 16: Inter-item correlation for recording strategies organised vs. disorganised..... | 76 |
| Table 17: Item-total statistics for recording strategies..... | 76 |
| Table 18: Cronbach's Alpha factor for recording strategies second investigation..... | 77 |
| Table 19: Mean strategy use..... | 77 |
| Table 20: Gender distribution of strategies..... | 80 |
| Table 21: T-test Gender and vocabulary strategies..... | 80 |
| Table 22: Cohens d factor..... | 80 |
| Table 23: Mean of vocabulary learning strategies and English speaking relatives..... | 81 |
| Table 24: Mean of Self-management strategies and school type..... | 81 |
| Table 25: Group statistics school type..... | 82 |

List of figures

| | |
|--|----|
| Figure 1: First method of adapting the GSL..... | 19 |
| Figure 2: Second method of adapting the GSL | 20 |
| Figure 3: Third method of adapting the GSL..... | 20 |
| Figure 4: Aspects of word knowledge (Nation 2013: 48)..... | 25 |
| Figure 5: Form (Nation 2013: 48)..... | 26 |
| Figure 6: Meaning (Nation 2013: 48) | 27 |
| Figure 7: Use (Nation 2013: 48) | 28 |
| Figure 8: Curry’s model for learning styles (1983: 8)..... | 30 |
| Figure 9: Gu’s and Nation’s model (2007: 82)..... | 33 |
| Figure 10: Taxonomy of vocabulary learning strategies used in this study..... | 35 |
| Figure 11: Discovery strategies | 42 |
| Figure 12: Consolidation strategies..... | 46 |
| Figure 13: Austrian Education system (Benedik et al. 2017: 17)..... | 56 |
| Figure 14: Distribution of students in school types (Benedik et al. 2017: 25)..... | 57 |
| Figure 15: Relationship between school type and tertiary education (Benedik et al. 2017: 63) | 57 |
| Figure 16: Development of academic vocabulary..... | 64 |
| Figure 17: Development of high-frequency vocabulary (1-2000) | 65 |
| Figure 18: Development of high-frequency vocabulary (2000-4000) | 66 |
| Figure 19 Development of low-frequency and technical vocabulary | 68 |
| Figure 20: Legend of proficiency levels | 68 |
| Figure 21: Academic vocabulary proficiency levels..... | 69 |
| Figure 22: High-frequency vocabulary (1-2000) proficiency levels | 70 |
| Figure 23: High-frequency vocabulary (2000-4000) proficiency levels | 70 |
| Figure 24: Divergent development of academic and high-frequency vocabulary | 71 |
| Figure 25: Information on gender, school type and experience abroad..... | 72 |
| Figure 26: Information on mother tongue, bilingualism and existence of English relatives | 73 |
| Figure 27: Gender and bilingualism in relation to high-frequency vocabulary (1-2000) 74 | |
| Figure 28: Academic vocabulary and planning strategies..... | 78 |
| Figure 29: Academic vocabulary and planning strategies..... | 79 |
| Figure 30: Legend of proficiency levels | 82 |

1. Introduction

In the last 15 to 20 years the interest in English as a lingua franca increased (Brosch 2015: 75), since English has developed into the main language used in several contexts, such as in education, business and popular culture (Seidelhofer, Breiteneder & Pitzl 2006: 3). Amongst other areas English is becoming the dominant language of research and publication in academia (Seidelhofer, Breiteneder & Pitzl 2006: 3). Therefore, academic writing in the English language has become a fundamental skill for university students around the globe who aim at making a contribution to academic discourse. As a result, Universities aim at training their students in academic writing to make them successful members of the academic society.

Since this tough goal of being a proficient academic writer challenges students and non-native scholars likewise, as reported by Paquot (2010: 2), several studies have been conducted to discover the truly academic features which are unique in academic discourse. One major discovery of applied linguistics has been that academic texts consist of a specific mixture of vocabulary types (Nation 2013: 289). With the help of corpus analysis programs, it was possible to differentiate academic vocabulary items which are highly frequent in academic discourse. Therefore, students wishing to become accomplished academic writers should try to enhance their knowledge of academic vocabulary amongst other skills. This paper is going to investigate with the help of a corpus analysis whether students of the English Department of the University of Vienna managed to reach this goal of increasing their academic vocabulary knowledge. Academic vocabulary items co-occur with other vocabulary types, namely technical vocabulary, high-frequency and low-frequency vocabulary. Therefore, in this research, the development of high-frequency vocabulary in relation to academic vocabulary will be examined as well. Low-frequency and technical vocabulary will be subsumed in one group, termed the rest, since they are not the primary focus of this research but give interesting insights into the distribution of vocabulary types in the texts. Furthermore, it will be analysed which vocabulary learning strategies the participants applied to reach their aim of enhancing academic vocabulary and if a relationship between strategy use and proficiency level can be found. Findings in these areas could make interesting contribution to existing literature and teaching and learning practice at Universities. While correlations between vocabulary learning strategies and learning success have

been investigated before, a novel perspective will be shed on the issue by investigating whether biographical information, such as school type attended, gender, mother tongue reported, experience abroad, existence of English speaking relatives or bilingualism have an influence on strategy choice and/or vocabulary development. The exact research questions examined are:

- 1) how academic vocabulary developed from the first seminar in academic writing to the last one;
- 2) how high-frequency, technical and low-frequency vocabulary developed in comparison to academic vocabulary;
- 3) whether students show varying proficiency levels in academic vocabulary;
- 4) whether a relationship can be detected for vocabulary learning strategies and biographical information;
- 5) whether a relationship can be detected for vocabulary development and vocabulary learning strategies used by the students;
- 6) whether a relationship can be detected for vocabulary development and biographical information supplied by the participants;
- 7) how the accuracy and formality of students developed during their studies.

The theoretical part of the paper preceding the practical part will thoroughly discuss all vocabulary types used in the academic texts already mentioned above. The paper further outlines the development of word lists of these vocabulary types. A theoretical chapter on corpus analysis programs will explain their use for this study and will open up possibilities for future investigations, since unfortunately the extent of this research paper only allows a narrow focus. This means that key areas for the investigation of academic vocabulary, such as collocations, are only going to be touched upon. However, opportunities for future research will be mentioned.

The second part of the theoretical input starts with an introduction of aspects of word knowledge necessary to be considered when attempting to learn a new word. This is followed by a broad distinction of learning style and learning strategies and goes on to discuss vocabulary learning strategies in particular. In this chapter, a novel taxonomy of vocabulary learning strategies combining existing approaches to taxonomies will be developed. This will be used as the basis for the empirical study. Furthermore, the influence that gender, bilingualism, experience abroad and school type visited had on vocabulary learning or education in general will be addressed.

2. Literature review

This chapter provides an overview of relevant literature to lay the foundation for the empirical part. Firstly, academic, technical, high- and low-frequency vocabulary needed in academic writing are defined. This distinction is essential since the empirical study investigates all vocabulary types mentioned. Secondly, this literature review provides an introduction to the development and qualitative features of the main word lists conducted, which will be compared and assessed according to their relevance for the practical application. Thirdly, the main corpus analysis programs in the field will be compared briefly to justify the later choice of methodology. In the next point, aspects of word knowledge needed to learn vocabulary sufficiently will be described. After a brief definition of learning style and learning strategy this will be linked to recent discussions about vocabulary learning strategies. In the course of the review of vocabulary learning strategies a new taxonomy of vocabulary learning strategies will be presented, based on recent approaches to strategy categorisation. This taxonomy will later form the basis for the practical examination of vocabulary learning strategies. Lastly, recent findings on vocabulary learning concerning the influence of the biographical background factors of gender, bilingualism, school type attended and experience abroad will be outlined.

2.1. Vocabulary in academic writing

Since the following study is concerned with an analysis of academic texts written by students, it appears reasonable to define the different vocabulary levels necessary to compose an academic piece of writing. Nation (2013: 289) reports that vocabulary use in academic texts has commonly been marked by a threefold differentiation into “general service [...], sub-technical [...] and technical vocabulary”. Paquot (2010: 9) elaborates on Nation’s distinction by paying attention to synonymous applications of terms when stating that academic vocabulary is identical to sub-technical vocabulary as well as discourse organising vocabulary, which are distinguished from core vocabulary and technical vocabulary. While both Nation and Paquot agree on a tripartite structure, Bruce (2011: 96-97) differentiates between high-frequency words, academic vocabulary, technical vocabulary and low- frequency words, therefore suggesting four categories. In contrast, Townsend and Kiernan (2015: 13) use academic vocabulary as an umbrella term for general academic vocabulary and discipline specific vocabulary, of which the latter can be equated to technical vocabulary in the other three models. For the purpose of this

paper a distinction between words of varying frequency, namely high- and low-frequency words, academic vocabulary and technical vocabulary, will be made. However, it is necessary to mention that the boundaries between these categories are rather fuzzy (Paquot 2010: 15).

2.1.1. High- and low-frequency words

High-frequency vocabulary is also known as general service vocabulary (West 1953: x), core vocabulary (Bussmann 1996: 49) or basic vocabulary (Nation 2013: 289). However, all three terms are typically associated with the most frequent 2000 words of the English language (Chung & Nation 2003: 103). Admittedly, these 2000 words are not always identical, since the language items might change over time and new words, such as *computer*, enter the group of highly frequent words. In spite of the varying terminology the current study will consistently use the term high-frequency words, acknowledging the importance of frequency of occurrence.

High-frequency words can be found in all contexts of written or spoken discourse (Paquot 2010: 10; Kwary 2011: 176). Chung and Nation (2003: 104) state that they cover 80 percent of the running words in an academic text and 90 percent in novels and conversations. Stubbs (1986: 104) describes high-frequency words in more detail by highlighting their pragmatic neutrality. He names five criteria which need to be fulfilled by a word to be counted as pragmatically neutral and therefore as a potential high-frequency word in addition to high-frequency of occurrence (Stubbs 1986: 104):

1. [High-frequency] words have a 'purely conceptual, cognitive, logical or propositional meaning, with no necessary attitudinal, emotional or evaluative connotations'.
2. They have no cultural or geographical associations.
3. They give no indication of the field of discourse from which a term is taken, i.e. its domain of experience and social settings.
4. They are also neutral with respect to tenor and mode of discourse: they are not restricted to formal or informal usage or to a specific medium of communication, e.g. written or spoken language.
5. They are used in preference to non-nuclear words in summarizing tasks.

Nation and Kyongho (1995: 35) along with Chung and Nation (2003: 104) and Paquot (2010: 10) further remark that high frequency words can be divided into function words and content words. Bruce (2015: 96) sharpens these statements by mentioning that high-frequency words include 176 different function words, while de Chazal (2014: 92) clarifies the nature of high-frequency content word by summarizing that high-frequency

language is useful, productive and generative. These attempts at a description of high-frequency vocabulary form the basis of this and other studies (Gu 2002: 38; Valipouri & Nassaji 2013: 248), since existing high-frequency lists used for investigation are mostly based on these assumptions.

Low-frequency words, in contrast to high-frequency words, can be described as having very specialized meanings, which explains their limited spectrum of occurrence (Bruce 2015: 96). They can be identified without great effort since they are items “that are not high-frequency words, not academic words and not technical words for a particular subject“ (Kwary 2011: 176). Therefore, depending on how clear the division between the other three categories is, the effort of low-frequency vocabulary identification varies.

2.1.2. Academic vocabulary

To define academic vocabulary several features need to be considered. Firstly, academic vocabulary is typical for academic texts since “academic vocabulary is common to a wide range of academic texts, and not so common in non-academic texts” according to Nation (2013: 291). His view is shared by other scholars such as Paquot (2010: 9) and Townsend and Kiernan (2015: 113), who state that “academic vocabulary words [...] are words that appear with much greater frequency in academic texts than in other types of texts”.

Secondly, another widely acknowledged feature of academic vocabulary is its “interdisciplinary nature” (Jordan 1997: 153). Coxhead (2000: 214) notes that “[a]cademic words are not highly salient in academic texts, as they are supportive of but not central to the topics of the texts in which they occur”. According to Nation (2013: 295) and Paquot (2010: 9) it is this property that differentiates academic vocabulary from technical vocabulary.

Lastly, Martin (1976: 91) and Nation (2013: 295) emphasize the educational value of academic vocabulary. Nation (2013: 292) states that academic vocabulary is generally less known to learners than the technical vocabulary of their specific fields. Therefore, language learners from all disciplines will benefit from an introduction to academic vocabulary. While Nation (2013: 292) stresses the general helpfulness of academic vocabulary for study goals, Martin (1976: 93) categorises academic words according to their specific use in academic discourse: academic vocabulary might be used in the research process, in analysis and evaluation. For Martin (1976: 93) the vocabulary of the

research process primarily consists of verbs and nouns which are used for “formulating, investigating, analysing, drawing conclusions and reporting results”, while the vocabulary of analysis “includes high frequency and two-word verbs needed in order to present information in an organized sequence”, such as “*consist of, result from and derive*”. The vocabulary of evaluation in turn is characterised by “adjectives and adverbs [that occur] in reviews, critiques and some reports, [such as] *exhaustive [...], controversial [or] coherent*” (Martin 1976: 95).

This study will not distinguish between academic vocabulary types used at specific stages of academic work. However, it is acknowledged that academic vocabulary is highly frequent and interdisciplinary in academic texts.

2.1.3. Technical vocabulary

Technical vocabulary is not solely relevant in English for specific purposes (ESP) contexts but for the analysis of academic texts in general (Kwary 2011: 175) since technical words cover 5 percent or more of the running text (Chung & Nation 2003: 104; Bruce 2011: 97).

A proper definition of technical vocabulary needs to focus on several aspects of technical vocabulary. Primarily, technical vocabulary can be described as “subject-specific” (Nation & Chung 2004: 252; Paquot 2010: 13; de Chazal 2014: 92), which means that “words are closely related to the content of a particular discipline” (Nation 2013: 303). Therefore, technical vocabulary clearly diverges from academic vocabulary concerning the degree of subject relatedness. However, technical terms are not necessarily unique to a single discipline but can belong to several related fields (Nation 2013: 304). For instance, *resistance* is a technical word in medicine meaning a “[n]atural or acquired ability to withstand disease, infection or attack by pests” but can also stand for “[an] electrical component considered with respect to its resistance” in the field of electronics (*Oxford English Online dictionary* 2010). Despite this over-disciplinary variation a technical word typically has a narrow and unique meaning connected to each discipline (Nation 2013: 304). Paquot (2010: 13) describes technical words as being characterized by “a semantic specialization, resistance to semantic change and absence of exact synonyms”. Therefore, technical terms are not only related to a certain discipline but are also irreplaceable in the language of a certain field.

The audiences for technical terms are “people working in a specialised field” (Chung & Nation 2003: 104). They are therefore part of a “specialist domain” included in a “system of subject knowledge” (Chung & Nation 2003: 252). This raises the question as to how technical vocabulary can be identified. Paquot (2010: 13) explains that frequency of occurrence, range and distribution play an elementary role in their selection. Paquot (2010: 13) and Nation and Kyongho (1995: 36) alike state that technical vocabulary occurs with high or moderate frequency in specialised texts. However, Nation (2013: 304) debates whether technical words can be high-, mid- or low-frequency words, explaining that highly frequent words can have a specialised connotation in a specific discipline and that academic words can also adopt a specialised meaning in certain contexts (Nation 2013: 304). Considering all three factors for choice, Chung and Nation (2003: 105) developed a four-scale item classification system for medical texts focused on anatomy. Raters were asked to classify words in medical texts with the help of the scale presented in Table 1.

Table 1: Identifying technical words (Chung & Nation 2003: 105)

| |
|---|
| <p>Step 1 Words such as function words that have a meaning that has no particular relationship with the field of anatomy, that is, words independent of the subject matter. Examples are: <i>the, is, between, it, by, 12, adjacent, amounts, common, commonly, directly, constantly, early, and especially.</i></p> |
| <p>Step 2 Words that have a meaning that is minimally related to the field of anatomy in that they describe the positions, movements, or features of the body. Examples are: <i>superior, part, forms, pairs, structures, surrounds, supports, associated, lodges, protects.</i></p> |
| <p>Step 3 Words that have a meaning that is closely related to the field of anatomy. They refer to parts, structures or functions of the body, such as the regions of the body and systems of the body. Such words are also used in general language. The words may have some restrictions of usage depending on the subject field. Examples are: <i>chest, trunk, neck, abdomen, ribs, breast, cage, cavity, shoulder, girdle, skin, muscles, wall, heart, lungs, organs, liver, bony, abdominal, breathing.</i> Words in this category may be technical terms in a specific field like anatomy and yet may occur with the same meaning in other fields and not be technical terms in those fields.</p> |
| <p>Step 4 Words that have a meaning specific to the field of anatomy and are not likely to be known in general language. They refer to structures and functions of the body. These words have clear restrictions of usage depending on the subject field. Examples are: <i>thorax, sternum, costal, vertebrae, pectoral, fascia, trachea, mammary, periosteum, hematopoietic, pectoralis, viscera, intervertebral, demifacets, pedicle.</i></p> |

The classification system was tested with the help of an inter-rater reliability check and altered according to the results. Vocabulary items falling under step one and two were not classified as technical words in medical contexts. Only words fitting into step three or four were defined as technical terms. While words at step three may still have polysemes which can be found in general language use, vocabulary items rated as step four were maybe generally known but with a “technical flavour” (Chung & Nation 2003: 105).

For other research areas a similar tool and procedure would be conceivable to identifying technical words. In this study technical vocabulary is investigated together with low-frequency vocabulary as one group. However, no distinction is made between the two types in this case, since the texts under investigation cover a broad range of topics and several disciplines. Therefore, a focus on technical vocabulary would probably not yield convincing results.

2.2. Word lists

In view of the diverse types of vocabulary necessary for successful academic writing, the question arises as to which words are most desirable to learn for which audience. To answer this, linguists have developed several types of word lists suitable for particular audiences. Bogaards and Laufer (2004: 3) note that such word lists are usually compiled on the basis of a needs analysis. They emphasise that “[in] any needs analysis it is important to decide whose needs are being investigated, and then to ensure that the investigation draws on data that is relevant to the people” (Bogaards & Laufer 2004: 3). This section will introduce the most relevant word lists for high-frequency vocabulary and academic vocabulary and will describe their intended audiences as well as their strengths and weaknesses. However, before focusing on specific examples, selection criteria for a high-quality word list will be introduced.

While prior to the rise of technology, content words were selected randomly from the most prestigious literary texts (de Chazal 2014: 103), nowadays words are chosen on the basis of frequency information (Nation & Waring 1997: 17). The more frequent a word is, the more likely students are going to be confronted with its meaning (Nation & Waring 1997: 17). Nation (2013: 14) additionally confirms that “frequency based studies show strikingly [...] that some words are much more useful than others”. However, this statement needs to be weakened for high-frequency vocabulary since high-frequency

items are only useful for beginners and need to be supplemented with other vocabulary types at higher levels. Nevertheless, also other types of vocabulary, such as academic or technical vocabulary, can be sorted by frequency, making them more likely to be encountered in specific fields. Therefore, one way of providing learners with suitable vocabulary is to give them frequency-based word lists for their specific learning purpose. Nevertheless, frequency based lists need to meet several criteria to be of high quality. Nation and Waring (1997: 18) list five distinct criteria, namely representativeness, limitations of frequency of occurrence, ways of counting, handling of collocations and range of information presented, worth considering when evaluating a frequency based word list. Their compilation has been complemented with remarks from other scholars:

2.2.1. Representativeness

Nation and Waring (1997: 18) remark that the corpus of texts from which the words are chosen needs to be representative. Considering several studies on representativeness six prerequisites for this criterion can be named which will be discussed in the following. According to Nation and Waring (1997: 18) both written and spoken texts should be considered, unless the research has a specific focus on written or oral production, such as this paper, which will focus on written texts. Moreover, both Nation and Waring (1997: 18) and Durrant (2016: 52) highlight the importance of including different text types. Durrant (2016: 52) additionally emphasizes the necessity to include texts from varying disciplines, since some words are much more common in particular fields than in others. Coxhead (2000: 216) adds that the texts should be equally spread between disciplines to make the study representative. She also suggests that the text corpus needs to be of a representative size to make valid judgements about frequency (Coxhead 2000: 216). Finally, Durrant (2016: 52) distinguishes between different proficiency levels of learners, who require a divergent selection of texts.

2.2.2. Frequency range and dispersion

Nation and Waring (1997: 18) advise that “[a] word should not become part of a [...] list merely because it occurs frequently [in one text]. It should occur frequently across a wide range of texts [in the corpus]”. For instance, if a corpus consisted of texts from Literature Studies, Physics and Economics, the range would indicate whether a word is part of all three or at least two of the fields of inquiry. Therefore, range indicates if a word is found

in most or all texts of a corpus. While Nation and Waring (1997: 18) are content with the measurement of range, Gardner and Davies (2014: 316) prefer the measurement of dispersion, which they consider to be “superior” to the range measure. They define dispersion as a measure which “shows how ‘evenly’ a word is spread across the corpus” (Gardner & Davies 2014: 316). Coming back to the example from above this means that dispersion would examine if the word was much more frequent in one of the disciplines where it is used. The degree of dispersion may vary from 0.01 to 1.00. If the dispersion is 0.01, the word is only frequent in a very small part of the corpus, while 1.00 indicates that the word is perfectly distributed over all parts of the corpus (Gardner & Davies 2014: 316). Summing up, range measures whether a word is found in most texts included from a certain discipline or genre, while dispersion measures how evenly distributed the word is across these disciplines or text types (Gardner & Davies 2014: 316).

2.2.3. Ways of counting

Word lists need to make use of consistent methods of counting words to be comparable. Nation (2013: 9) broadly distinguishes between types and tokens. If the tokens of a text are counted, repeated words are counted each time, while if types are noted, a word is only counted once (Nation 2013: 9). For instance, looking at the previous sentence the word *counted* would account for two tokens but only one type.

Some words are closely related to each other, such as the singular or plural forms of a word. To account for such relationships two counting systems are used, namely lemmas or word families. Nation (2013: 10) defines a lemma as follows: “[a] lemma consists of a headword and its inflected forms and reduced forms”. This means that in English a lemma can contain plural, third person singular present tense, past tense, present participle, past participle, comparative, superlative and possessive inflections depending on the word class (Nation 2013: 10). A word family in contrast “consists of a headword, its inflected forms” and immediately derived forms, such as the headword plus regular affixes (Nation 2013: 10). Therefore, word families incorporate more variations under one term. Both systems are legitimate but have their flaws, which will be discussed in more detail at a later stage of this section.

2.2.4. Collocations

Native speakers of a language tend to store vocabulary as prefixed chunks instead of individual words (Wray 2000: 464; 2002: 6; 2004; Durrant & Schmitt 2010: 163; Nation 2013: 482). While Wray (2000: 468; 2002: 148) claims that language learners rely on individual words instead of chunks, the studies of Durrant and Schmitt (2009; 2010) indicate that adult second language learners also use chunking. However, in contrast to native speakers they rely heavily on the use of highly frequent collocations, which they tend to use repetitively, while natives use more strongly associated low-frequency collocations (Durrant & Schmitt 2009: 157).

To ensure that studies are comparable the question needs to be asked as to how collocations can be defined. Lehecka (2015: 4) distinguishes three types of definitions for collocations. The Firthian tradition (see Firth 1957; 1968) states that “collocations are statistically significant co-occurrences of two or more words regardless of the meaning of these word combinations” (Lehecka 2015: 4). Studies coming from a phraseological background define collocations as “a word combination that has been lexicalized to at least some extent” (Lehecka 2015: 4). Wray (2002: 9) describes this type of collocation as “a sequence, continuous or discontinuous, of words or other elements which is [...] stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar”. Computational linguistics has given rise to a third definition of collocations, namely that they are “a word combination with idiosyncratic semantic or syntactic properties” (Lehecka 2015: 4). Evert (2009: 1214) has named these three types empirical collocations, lexical collocations and multiword expressions.

Moon (1997: 43) defines a multiword unit as “a vocabulary item which consists of a sequence of two or more words. This sequence of words semantically and/or syntactically forms a meaningful and inseparable unit”. Nation (2013: 480) extends her definition by naming four overlapping characteristics of multiword units. Firstly, he highlights the flexible nature of multiword units by stating that the word order of a collocate might change or that additional words to the main collocates can be altered (Nation 2013: 480). Secondly, he mentions the special morphology and semantics of a multiword unit. While in general the morphological and semantic use of the individual words of a multiword unit is “consistent with [the] typical grammatical and semantic use outside any particular

multiword unit”, they might influence the morphological form of their collocates (Nation 2013: 480). However, Nation (2013: 480) extends in a third point that semantically certain types of multiword units can mean more than the individual collocates do. They form their own extended semantic meaning (Nation 2013: 480). Lastly, he notes that multiword units similar to normal single words follow a communicative purpose in interaction (Nation 2013: 480)

Moon (1997: 44) names three characteristics which distinguish holistic multi-words from other collocational formations. Firstly, multi-words can vary in their degree of institutionalisation in a language system, which means in a narrower sense that “the degree to which a multi-word is conventionalised in the language” might change (Moon 1997: 44). Secondly, multi-words can have different degrees of fixedness, which means that some strings are frozen, while others remain more flexible. For instance, the idiom *beat around the bush* is strongly fossilised, since the word *bush* cannot be exchanged for any other word without a loss of meaning. Thirdly, Moon (1997: 44) mentions non-compositionality, which describes the degree to which a unit can be interpreted through the meaning of its individual words. Referring to the example from above, the idiom *to beat around the bush* is non-compositional, since the semantic meanings of the individual words do not add up to the meaning these words have in combination, namely to avoid the main topic. Nation (2013: 83) extends this list of features by indicating that collocations and therefore also multiword units can vary in the number of words involved, in the type of words involved, as well as in the closeness of the collocates. Moon (1997: 44-47) distinguishes between compounds, phrasal verbs, idioms, fixed phrases and prefabs, which vary in the degree of institutionalisation, fixedness and non-compositionality. In the following definitions the five categories distinguished by Moon (1997: 44-47) are briefly summarized:

- According to the concise Oxford Dictionary of linguistics (2014) a compound is “[a] word formed from two or more units that are themselves words or forms of words”. Possible examples are *blackboard* or *wildflower*. Compounds can be further distinguished into noun, verb and adjective combinations and are the most extensive group of multiword units (Moon 1997: 44).
- Phrasal verbs “are combinations of verbs and adverbial or prepositional particles” (Moon 1997: 48).

- An idiom is “[a] set expression in which two or more words are syntactically related, but with a meaning like that of a single lexical unit (*The concise Oxford Dictionary of linguistics* 2014).
- The term fixed phrases is used as an umbrella term for words which are not included in the previous categories or in the last one (Moon 1997: 49). While most fixed phrases are strongly institutionalised and strongly fixed, their compositionality may vary (Moon 1997: 49).
- “Prefabs are preconstructed phrases, phraseological chunks, stereotyped collocations or semi-fixed strings which are tied to discoursal situations and which form structuring devices” (Moon 1997: 49).

As already mentioned in the introduction this paper will not investigate any types of collocations. However, this does not mean that it does not acknowledge the essential role of collocations for second learners’ language use. Therefore, an investigation of collocations would be a crucial addition to this study and should be considered for future research. Especially multi-word units pose an interesting field of research, since there have not been too many investigations of the development of multi-word units in academic contexts.

2.2.5. Range of information

Nation and Waring (1997: 19) state that for educational purposes word lists need to include information about “the forms and parts of speech included in a word family” and how frequent the individual words occur. Furthermore, it would be crucial to include information on the different semantic shades that a word can express and which collocations it forms with which frequency. To make the words in the list useful for communicative purposes, social aspects need to be considered, such as the degree of formality of a word and whether it is considered to be polite or impolite or culturally biased. Modern lists, such as the Academic Vocabulary List by Gardner and Davies (2014) attempt to pay attention to these pedagogical implications. Other scholars even offer training tools for the words included in their word lists, such as Browne, Culligan and Phillips (2013) did for their New Academic Word List. Nevertheless, this criterion is not that important for the purpose of this study, since word lists are only used for investigation and not for didactic purposes.

2.2.6. Academic and high-frequency word lists

After this introduction to the criteria for qualitative word lists, the most relevant high-frequency and academic word lists will be discussed.

The most frequently used, although recently criticised, list of high-frequency words to date is the General Service List (GSL), which was produced by West in 1953 (Nation & Kyongho 1995: 35). He compiled a corpus of five million running words and ordered them according to frequency. However, he emphasizes that five other features alongside frequency need to be considered to make a selection of words, namely difficulty of learning, necessity, cover, stylistics and intensive or emotional words, which are explained below (West 1953: ix-x). Difficulty of learning has also been considered by other scholars, such as Nation (2013: 45). He names this concept the learning burden, defining it as “the amount of effort required to learn [a word]” (Nation 2013: 45). Possible influences on the perceived learning burden for certain words will be discussed in more detail in section 2.4.1. in relation to aspects of word knowledge. Considering necessity, West argues that some low frequency words need to be included in the list as well since they are the only umbrella term for many words (West 1953: ix). As an example, he names the word *preserve*, which covers the entire semantic range of canning, bottling or freezing food (West 1953: ix). The notion of cover excludes highly frequent words which are unnecessary though (West 1953: x). Concerning emotional words, he claims that all words used as intensifiers in English, such as *simply useless* expressing uselessness and annoyance, can be left out by beginners, since emotion is tied to the mother tongue and not to the second language learned (West 1953: x). This argument appears to be rather weak, since learners clearly have emotional attitudes towards and in their L2 as can be seen in language attitude studies (Hiraga 2005; Chevalier 2014).

Additional to this judgement West’s GSL has been criticised due to its age (Gardner & Davies 2013: 308) and size (Browne, Culligan & Phillips 2013). Nation and Kyongho (1995: 36) recommend lists compiled by Thorndike and Lorge (1944), Carroll et al. (1971) and Francis and Kucera (1982) for learners as supplementation to the GSL. However, new attempts to produce high-frequency lists have mostly been discipline specific, since large parts of the research field of corpus linguistics hold the opinion that learners should focus on their specific field of interest and not on general service vocabulary (Hyland & Tse 2007: 235; Chen & Ge 2007: 513). However, other studies have provided convincing evidence that neutral vocabulary not linked to any particular

discipline exists and that it is equally beneficial for all types of learners (Gardner & Davies 2013: 306; Durrant 2016: 60). Therefore, a renewed version of the GSL has been produced for this study as explained in more detail at the end of this section. This adaptation of the GSL will be used as baseword list in the empirical part of the study.

In his monograph on academic vocabulary, Hirsh (2010) offers a historical overview of the development of academic word lists, which he describes, compares and partly evaluates. He states that the first academic word lists to become highly influential were based on four studies conducted in the 1970s. Champion and Elley (1971) developed an academic vocabulary list grouped in lemmas for students attempting a university entrance exam in New Zealand. They examined diverse academic texts, such as textbooks, lectures or examination papers and came up with a list of the 500 most common academic words and a total list with 3200 academic words (Hirsh 2010: 17). Like Champion and Elley (1971), Ghadessy (1979) prepared two lemma-based word lists for academic purposes, considering twenty science textbooks for his investigations (Hirsh 2010: 18). The first list was an extended version of the second list, which was compiled for a development program of first year university students (Hirsh 2010: 18). Praninska (1972) and Lynn (1973) in contrast grouped their findings into word families (Hirsh 2010: 20-21). Praninska's list covered 507 word families, while Lynn's included only 197 word families (Hirsh 2010: 17-18). Like Champion and Elley (1971), Praninska (1972) used a broad range of academic disciplines for her study, while Lynn (1973) is in line with Ghadessy (1979) focusing on one particular subject area, namely science (Hirsh 2010: 17-18).

Since these four lists had a relatively small impact individually, they were combined by Xue and Nation (1984) to form the University Word List, which contained 737 word families covering 8.5 percent of the running words of any academic text (Hirsh 2010: 25). In 2000 Coxhead (2000: 213) made a fresh attempt at compiling an academic word list, which would prove to become the most influential academic word list for the years to come (Coxhead 2011: 360). Her corpus consists of 414 academic texts from 28 disciplines, which make up 70,377 word families (Hirsh 2010: 26). One can see that the size of the corpora investigated increased over time with Praninska's (1972) study examining a corpus of 247,466 words and Coxhead's lists already investigating 3,513,330 running words (Coxhead 2000: 217). With the size of the sample the amount of detailed information about lexical items increased as well (Coxhead 2000: 216). Coxhead's (2000: 222) list is built on West's (1953) GSL, which means that it excludes all words on this list.

Similar to the GSL it is organised into word families, which is “supported by evidence suggesting that word families are an important unit in the mental lexicon” (Coxhead 2000: 218). The Academic Word List (AWL) outperforms the University Word List with regard to coverage of running words by accounting for 10 percent of all running words in her corpus (Coxhead 2000: 226). Together with West’s (1953) GSL a coverage of 86 percent of all running words was achieved (Coxhead 2000: 222). Therefore, the AWL appears to be a progress compared to previous word lists.

Nevertheless, further years of investigation gave rise to criticism, which resulted in the development of the Academic Vocabulary List (AVL) by Gardner and Davies in 2014. They argued that a new academic word list is necessary since Coxhead’s methodology had some flaws (Gardner & Davies 2014: 307). Firstly, they criticised her organisation of words into word families, since in “an extensive word family [members might] not share the same core meaning” (Gardner & Davies 2014: 307). As an example, they name *react* and *reactionary*. While the core meaning of *react* is to respond, *reactionary* means “strongly opposed to social and political change” (Gardner & Davies 2014: 307). Therefore, words subsumed under this word family have rather divergent meanings. They argue that parts of this divergence could be reduced if word families took account of grammatical parts of speech, since then for instance the noun *proceed* and the verb *proceed* would then be identified as separate units (Gardner & Davies 2014: 308). As a solution, they propose that words should be grouped in lemmas instead of word families, which they define as “words with a common stem, related by inflection only, and coming from the same part of speech” (Gardner & Davies 2014: 308), thereby, narrowing Nation’s (2013: 10) definition given above, since they do not include reduced forms of words, such as *I’ve*. This would enable the learner to realise the semantic difference between words in a more effective manner, since for example the noun *procedure* meaning technique would be counted separately from *proceedings* standing for records or minutes (Gardner & Davies 2014: 308). Secondly, they complain that Coxhead’s list is based on West’s GSL, which is relatively outdated (Gardner & Davies 2014: 308). Based on this assumption they claim that the AWL contains high-frequency words which were just not included in the GSL due to its age, such as *computer* or *internet*. However, these words would actually belong in the GSL and are not necessarily academic (Gardner & Davies 2014: 309). Another major and well-founded point of criticism adds that the GSL contains high-frequency words which happen to be used in academic texts as well, such as “*company, interest, business*

and market" (Gardner & Davies 2014: 309), which means that when using the AWL alone some crucial academic words are missing and when combining it with the GSL some academic words are counted as high-frequency words (Gardner & Davies 2014: 309-310).

Based on this criticism Gardner and Davies (2014: 312) developed their AVL, which is based on lemmas instead of word families. Their corpus consists of 120 million written words, which means that it is nearly 35 times larger than the corpus of the AWL (Gardner & Davies 2014: 313). It accounts for an even balance between nine academic disciplines (Education, Humanities, History, Social science, Philosophy & Religion, Law & Political Science, Medicine & Health, Business & Finance) and is focused on academic journals in combination with "academically oriented magazines" (Gardner & Davies 2014: 313- 314). The AWL in contrast included texts belonging to four academic disciplines, namely arts, commerce, law and science, which were again divided into 28 subject areas (Coxhead 2000: 216). To directly compare the AVL with the AWL, Gardner and Davies transformed their lemma based word list into word families, which showed that the AVL has nearly twice the coverage of running words as the AWL and shows similar coverage across all disciplines, indicating that it is a purely academic list with no technical bias (Gardner & Davies 2014: 322-323). Nevertheless, Durrant (2016: 50) points out that the value of the list lies more in receptive knowledge than in productive knowledge, since Gardner and Davies focus on texts written by professionals for an academic audience. Durrant (2016: 50) claims that this focus is due to the exclusion of student texts in the corpus, which suggests that he is implying that a corpus based on student writing would be more suitable to create an academic word list used to examine the academic vocabulary development of students. However, it might be argued that successful student writing should assimilate as far as possible to professional academic writing. Therefore, it appears reasonable to use authentic academic texts as models for students' reading and writing. Nevertheless, a more representative corpus of text, including argumentative essays, literature reviews or papers, which students might be required to write more often than research articles might be beneficial as long as the accuracy was given. In general, the AVL appears to be the most carefully considered academic word list at the moment and will therefore be used in the current study. However, in this research the AVL grouped into word families will be used and not into lemmas, since it will be employed in combination with the GSL, which is compiled in word families. While the points of criticism uttered against word families are valid they are not as substantial for the purpose of this paper,

since the list will not be used for pedagogical purposes but for a quantitative analysis. Therefore, the grammatical distinction provided by lemmas is not as relevant as for learning or teaching these lists.

Despite the points of criticism mentioned above, it can be stated that the GSL still provides a discipline-unspecific compilation of high-frequency words and therefore serves as a good basis for the investigation, since the texts to be analysed cover a broad spectrum of themes. More importantly, the GSL was used by the majority of conducted studies in the past, which means that using it as a foundation will guarantee a greater comparability of results than using a more recent list. However, it is undeniable that the GSL is outdated to some extent and lacks accuracy concerning the categorisation of words. To react to these flaws the GSL has been altered and will be used together with the AVL instead of the AWL. While the AVL will be used to investigate the academic vocabulary, the revised GSL will be utilized for the examination of the high-frequency vocabulary. Due to the more recent nature of the words collected in the AVL in comparison to the AWL words listed in the altered version of the GSL are more current than in the original GSL. Additionally, this combination ensures that the GSL was cleared of all high-frequency academic words by deleting all words from the GSL which also occurred in the AVL. This means that when examining the academic and high-frequency vocabulary of the participants the percentages of the vocabulary types used in the texts will be more accurate since no more academic words are grouped as high-frequency words just because they were listed on the GSL.

In Figure 1 below three columns can be seen. The column on the right is the opened AVL, while the column in the middle is the opened original GSL. The open program on the left is the corpus analysis program AntProfiler. As can be seen AntProfiler reports errors concerning certain words, such as *disagree*, *disagreeable* and *disagreement*. The error message reported means that the words reported have been found in the GSL and in the AVL. Therefore, these words are academic words according to the AVL but high-frequency words according to the GSL. The program is not sure, which vocabulary group it should add the words to and reports the error. Consequently, these duplicate entries need to be deleted from the GSL to get reliable results. This was now attempted with three different methods.

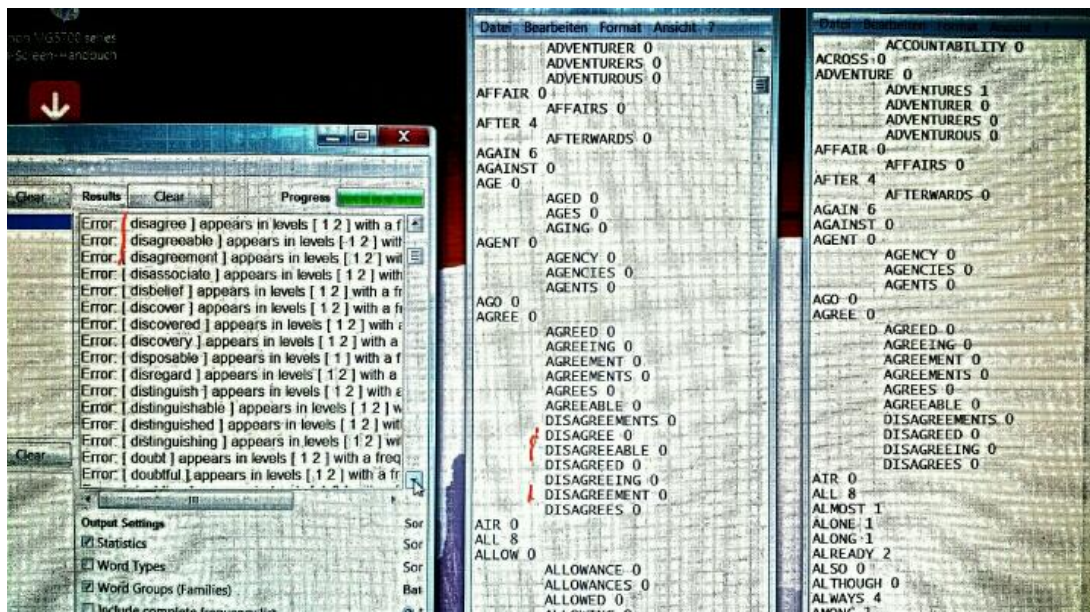


Figure 1: First method of adapting the GSL

If the words which were found twice (once on the AVL and once on the GSL) were only part of the word family and had no effect on the stem of the word family (words indented to the left), such as can be seen in Figure 1 above with *disagree*, *disagreeable* and *disagreement* from the family *agree* highlighted in red, these words were simply erased from the GSL (middle column), while they remained in the AVL. Therefore, *disagree*, *disagreeable* and *disagreement* will be counted as academic words in the next calculation.

However, in rare cases also the stem of the word family in the GSL was affected as in Figure 2 below. Here AntProfiler not only an error for words in the word family, such as *choosing* or *chosen* but also indicates that the word stem *choose* of the word family has been found twice. In this case two possible options were available. If another word of the word family was listed as possible stem of the family on the website of the British Council (<https://www.learnenglish.org.uk/wff/index.html>) this word was substituted for the previous stem. This was the case for *choose* in Figure 2 below. According to the British Council *choice* could also be a possible stem of this family. Therefore, *choose* was exchanged with *choice*. This substitution prevented the whole word family from being erased, since a word family cannot exist without a stem, so that words found only in the GSL, such as *choices* or *chooses*, would have been lost.

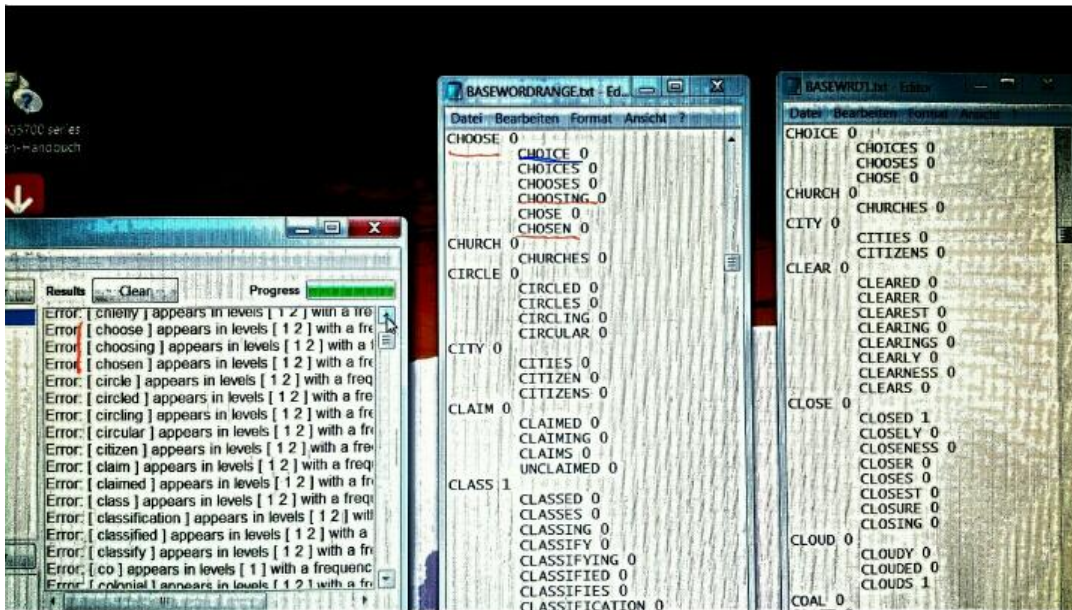


Figure 2: Second method of adapting the GSL

However, if there was no other possible stem acknowledged by the British Council or all possible stems were affected the whole word family was deleted, as can be seen with the word family *able* in Figure 3.

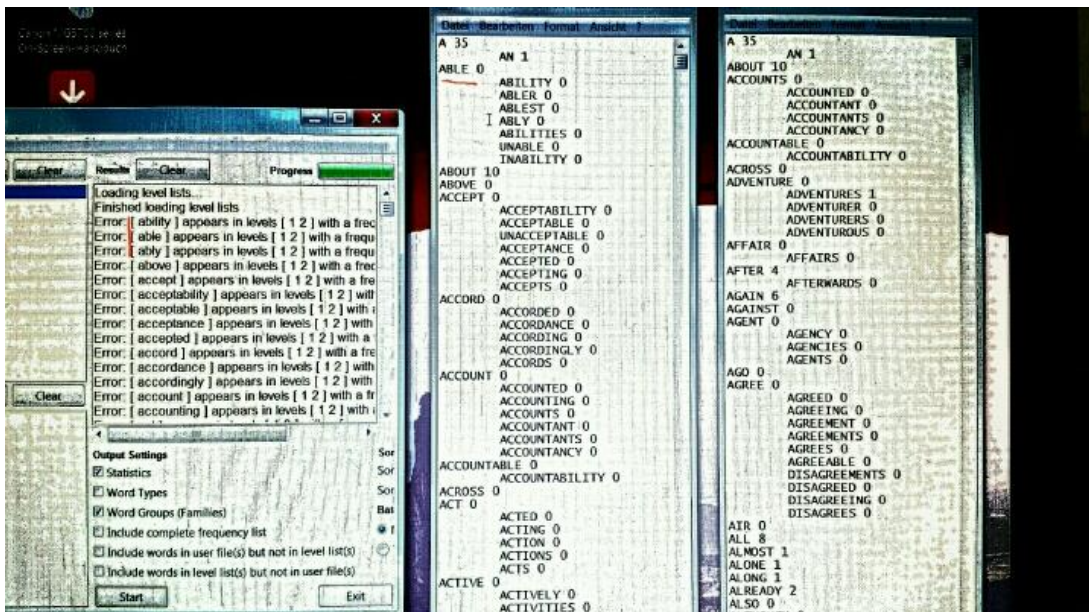


Figure 3: Third method of adapting the GSL

These alterations ensure that the version of the GSL used in this study definitely only includes neutral high-frequency vocabulary and no academic words. Therefore, results are going to be more accurate than with the unchanged GSL.

2.3. Corpus analysis programs

Computer programs have become a valuable and time-saving tool in the last years for investigations in corpus linguistics. Since a broad range of programs enabling researchers to analyse different aspects of language have been developed, the following section presents a short description and comparison of the programs considered for the corpus analysis of this study. Firstly, the programs Range (Nation 2005) and AntProfiler (Anthony 2014) will be introduced. These two programs were both suitable for the empirical part of this paper and can be used for similar examinations. Secondly, suggestions on how the research conducted in this paper could be further expanded will be given by introducing the programs AntConc (Anthony 2016), D-Tools (Meara & Miralpeix 2015) and V-Size (Meara & Miralpeix). However, these programs were not used in this study, since this vocabulary investigation extends the scope of this paper. Nonetheless, they are highly recommendable for future investigations.

The program Range (Nation 2005) examined in this paper is the freeware version by Nation from 2005. Range (Nation 2005) is a corpus analysis program which allows an analysis of up to thirty-two texts simultaneously. It provides the user optionally with different information. Firstly, it creates a distribution diagram of the range of a certain word in the texts (Nation 2005: 1). Therefore, if a teacher wishes to find out in how many texts a certain word can be found, as to find out if a certain word which had been studied previously has entered the students' active vocabulary, this tool can be useful. Moreover, the same can be done for headwords and whole word families. Additionally, the exact number of occurrences of the word is given. Another tool in the program enables users to find out how many words from a particular word list are given in a text, while another one can be used to create word lists based on frequency and range from an own corpus of texts (Nation 2005: 2). The program provides users with three included word lists based on the GSL and the AWL (Nation 2005: 6). The GSL has been divided into two separate lists, the first one including the 2000 most frequent words and the second one including the 2000-4000 most frequent words. This division was adopted for the altered GSL in this study, which revealed interesting results. Additionally, other word lists can also be uploaded to the program. These need to be coded in a certain format for the program to be able to read the list (Nation 2005: 7). If requested, the program also highlights words in the original text according to which list they belong to.

AntProfiler (Anthony 2014) is a relatively similar program developed in 2014 by Anthony in consultation with Nation, the creator of Range, who gave feedback on the program (Anthony 2014: 7). Similar to Range (Nation 2005) AntProfiler (Anthony 2014) has a vocabulary profile tool which “allows you to generate vocabulary statistics and frequency information about a corpus of texts loaded into the program” (Anthony 2014: 2). Additionally, it compares the texts with word lists. The same word lists are available as in Range (Nation 2005) but again other lists can be uploaded and the existing lists deleted (Anthony 2014: 3). The second tool available on AntProfiler (Anthony 2014) is the File Viewer and Editor tool. Inspired by Range (Nation 2005) this tool highlights the words in a running text according to the different vocabulary levels that the words belong to with the help of colour coding. However, in this program it is possible to immediately edit the file and to see the effect of the change directly in the program (Anthony 2014: 4). For instance, this can be used successfully to simplify a text for graded reading in school. For the purpose of this study the corpus analysis program AntProfiler (Anthony 2014) was used to compute statistical and frequency information about the coverage of word lists in the academic texts. Both programs, Range (Nation 2005) and AntProfiler (Anthony 2014), are rather equivalent in the quality and extent of the tools available. However, since AntProfiler (Anthony 2014) was developed later and in cooperation with Paul Nation it appears reasonable to see AntProfiler (Anthony 2014) as an updated version of Range. Moreover, AntProfiler (Anthony 2014) seems easier to use and has a less complex interface.

Unfortunately, the extent of this paper will not allow a consideration of all relevant aspects of vocabulary in academic writing. However, in the following the programs AntConc (Anthony 2016), WordSmith (Lexical Analysis Software & Oxford University Press 2017), V-Size (Meara & Miralpeix) and D-Tools (Meara & Miralpeix 2015) will be introduced, which could be used for a complementary research project to intensify and broaden the research scope.

AntConc (Anthony 2016), another corpus analysis program, could be used to investigate collocates in texts more closely (Anthony 2016: 6). However, each collocate needs to be searched for individually and typed in manually, which means that the collocations investigated need to be chosen and that the analysis is enormously time consuming or only small numbers of collocations can be examined. For a more automatic and time-saving investigation of collocations, the program WordSmith (Lexical Analysis Software

& Oxford University Press 2017) can be recommended. It allows users to generate a list of collocates which exist in a corpus of texts and it is able to compare a collocation list with the corpus. AntConc (Anthony 2016) is a useful program to examine specifically interesting collocations or when investigating different semantic uses of a word. In the concordance plot tool of AntConc (Anthony 2016), it can be seen if a word or collocate is more frequent in a certain part of a text and if it is evenly distributed (Anthony 2016: 3). This can be useful to investigate the textual functions of particular words.

Another interesting point for an investigation of vocabulary in academic texts is the lexical richness of a text. The programs D-Tools (Meara & Miralpeix 2015) and V_Size (Meara & Miralpeix) both calculate lexical richness (Meara & Miralpeix 2015: 1). However, D-Tools (Meara & Miralpeix 2015) works on the basis of a type-token ratio, while V_Size (Meara & Miralpeix) is operating with Zipf's law. Type-token ratios have been criticized to be prone to errors, since they depend on the length of texts (Meara & Miralpeix 2015: 1-2). Malvern and Richards (2004) offer a solution to these problems by arguing that the complexity of a text is better represented by individual repeated measurements instead of one large one. These type token ratios describing text samples of different sizes can be summarized by the parameter D (for more information see Malvern & Richards 2004). This solution is taken up and integrated in D-Tools.

V_Size on the contrary is based on the assumptions by Zipf, who developed a law which says that there is a direct proportional relationship between the number of times that a word occurs in a corpus and its rank order on the frequency list of the same corpus (Popescu, Altmann & Köhler 2010: 713). For instance, if the most frequent word of a corpus occurs 50 times in the texts, the word with the frequency rank 50 will occur 1 time. Therefore, the product of the rank on the frequency list and the times of occurrence is nearly identical according to Zipf's law.

V_Size uses this idea by taking several individual steps to generate a lexical richness profile (V_Size Manual: 2). Firstly, the program generates a word list for the texts uploaded to the program. Secondly, this word list is compared to a frequency dictionary built into the program. According to this, the program groups the words in the texts into several bands. Band A incorporates all words belonging to the most frequent 500 words, Band B includes the next 500-1000 most frequent words and all other bands work with steps of 500 as well. These numbers are then converted into percentages. In a next step

the program compares the results for these specific texts to theoretical curves stored in the program and tries to find the closest match. With this information a profile of the language user is estimated giving the rough vocabulary size of the text's author.

It has been outlined that both programs use completely different theoretical considerations as a basis for their computation of the lexical richness of a text. Therefore, it would be definitely interesting to use both programs for the same corpus of texts and to compare the results. Moreover, this might also be the safest option to generate correct results, since both theories appear to have their flaws. Zipf's law appears to be a relatively weak foundation on which to build estimates of the complete vocabulary size of a speaker. V_Size might calculate the vocabulary size of a speaker that was exploited in a particular text but this does not allow conclusions about the overall vocabulary size of a speaker, since the text analysed might be a particularly weak or strong text of the author. Therefore, regardless of the interesting results about lexical richness computed by such programs, results need to be treated with caution and should not be over-generalised.

2.4. Vocabulary learning

In the following the literature on vocabulary learning is reviewed. Firstly, the aspects will be investigated which learners need to consider to be able to be confident of knowing a word. Then the terms learning style and learning strategy are explained briefly, followed by a closer investigation of language learning strategies. Lastly, a conglomeration of several taxonomies of vocabulary learning strategies will be introduced as a basis for the empirical part of the paper. The individual categories and strategies will be described individually on the basis of existing literature.

2.4.1. Aspects of word knowledge

Figure 4 was designed according to Nation's (2013: 48) suggestions on what knowing a word entails.

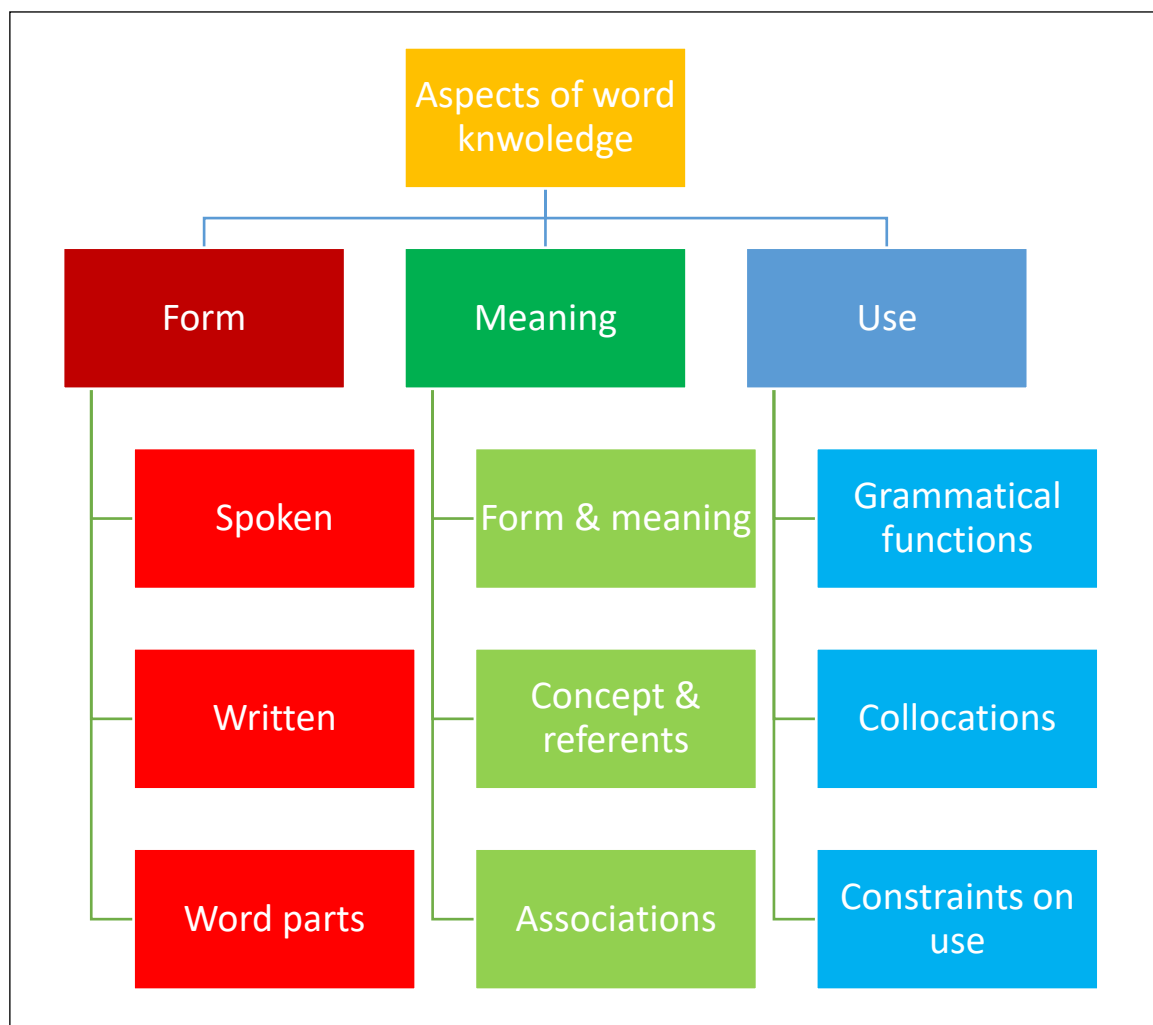


Figure 4: Aspects of word knowledge (Nation 2013: 48)

These individual aspects have to be considered with the learner's knowledge about two continuums. Nation (2013: 44) states that all aspects of word knowledge can be located on a continuum between item knowledge and system knowledge. For instance, the word *loved* requires item knowledge to grasp the semantic meaning of the word. However, to understand the full meaning of the word, knowledge about the system of language is necessary, such as, in this case, the formation of the regular past tense.

A second continuum can be defined, namely the distinction between receptive and productive knowledge. Palmer (1921: 118) has made the distinction between receptive skills, listening and reading, and productive skills, speaking and writing, which has been

preserved until today. Nevertheless, this categorisation is not as clear as suggested, since some tasks require a mixture of abilities. Nation (2013: 47) has utilized this principle successfully for the description of receptive and productive vocabulary usage:

Essentially, receptive vocabulary use involves perceiving the form of a word while listening or reading and retrieving its meaning. Productive vocabulary use involves wanting to express a meaning through speaking or writing and retrieving and producing the appropriate spoken and written form.

Therefore, the following aspects of word knowledge need to be viewed as being part of these continuums. To know a word a learner needs to be able to retrieve information about the form, meaning and use of a word.

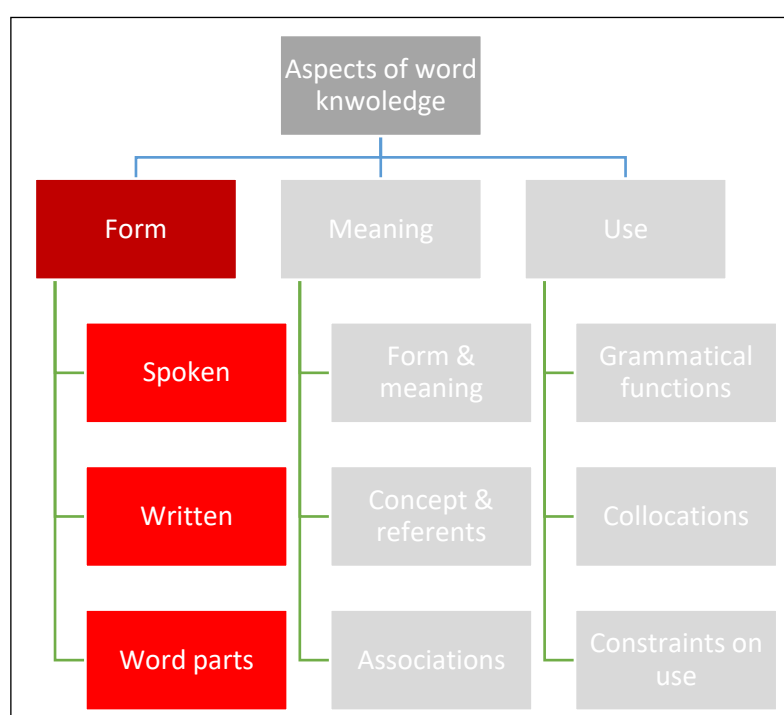


Figure 5: Form (Nation 2013: 48)

To be informed about the form of a word language users need to be able to recognise the spoken word when heard and to produce it correctly. The success of this production is heavily influenced by the pronounceability of a word, which partly depends on the similarities of sound and intonation between the L1 and L2 (Ellis & Beaton 1993: 559). Moreover, learners remember a word more easily if they can divide it into phonological units (Gathercole & Baddeley 1989: 200). The same accounts for the written form of a word. A strong connection has been found between pronunciation and spelling patterns (Bradley & Huxford 1994:). This means that learners remember a system of how sounds are reflected in spelling. Bradley & Huxford (1994:) report in their study that participants

who attended training in the matching of sounds and letters had better scores in accuracy of spelling and the effect persisted for years. Additionally learners need to be able to divide words into word parts. If word parts and word formation rules have been internalized, whole word families can be remembered (Nation 2013: 72).

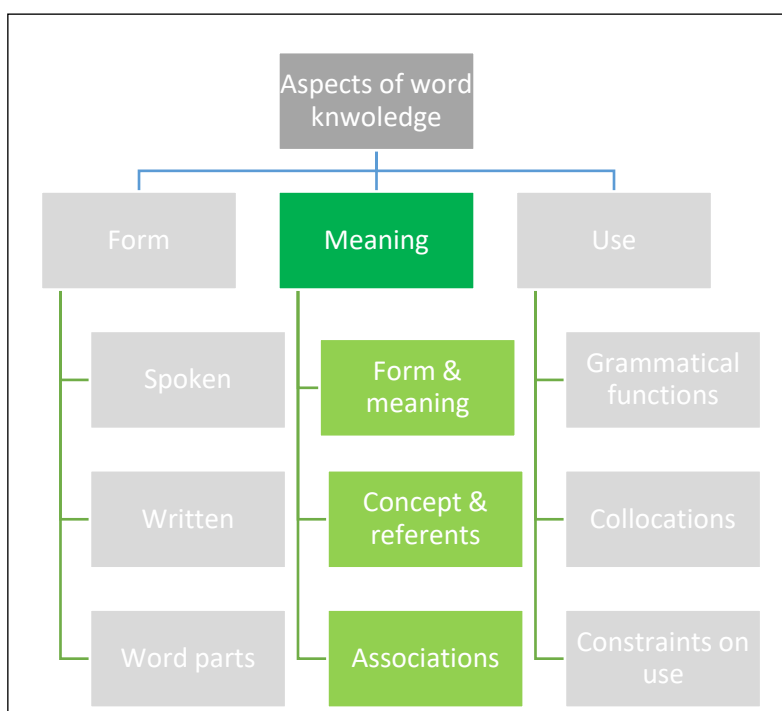


Figure 6: Meaning (Nation 2013: 48)

In a next step learners need to establish a connection between form and meaning. With repeated occurrence the form of a word and the semantic concept behind a word are strengthened (Nation 2013: 73). Another important issue especially when language use becomes more complex is that words do not only have one but several meanings. Taking a look at a random word in the Oxford paperback Thesaurus (Waite 2006) validates this assumption. For example, the verb *collapse* is used with the meanings “fall down (the roof collapsed), faint (he collapsed last night), to go to pieces (she collapsed in tears) and fail (peace talks collapsed)” (Waite 2006: 133). Nagy (1997) presents two approaches that learners might take to select the appropriate meaning of a word, namely sense selection and reference specification. Sense selection requires that the learner has stored each individual semantic shade of a word and has to select the correct one by considering the context. Reference specification only involves remembering an umbrella concept, such as a two-pronged shape for different types of forks, and the learner has to work out which specific meaning is suitable with the help of context. Both approaches can be used to acquire the meaning shapes of a word.

Learners should also know about the linguistic relationships that words can have, such as synonyms or homonyms and semantic relations between word types. Nouns, verbs and adjectives have their own internal organisation systems worth considering (Nation 2013: 79). Miller and Fellbaum (1991) investigated the internal structure of noun groups. They concluded that nouns can be ordered according to hierarchies in the shape of semantic trees. They claim that no more than 26 beginning points are needed to cover all nouns of the English language (Miller & Fellbaum 1991: 204-5). Learners might benefit from this knowledge, since they can organise words around concepts in semantic maps.

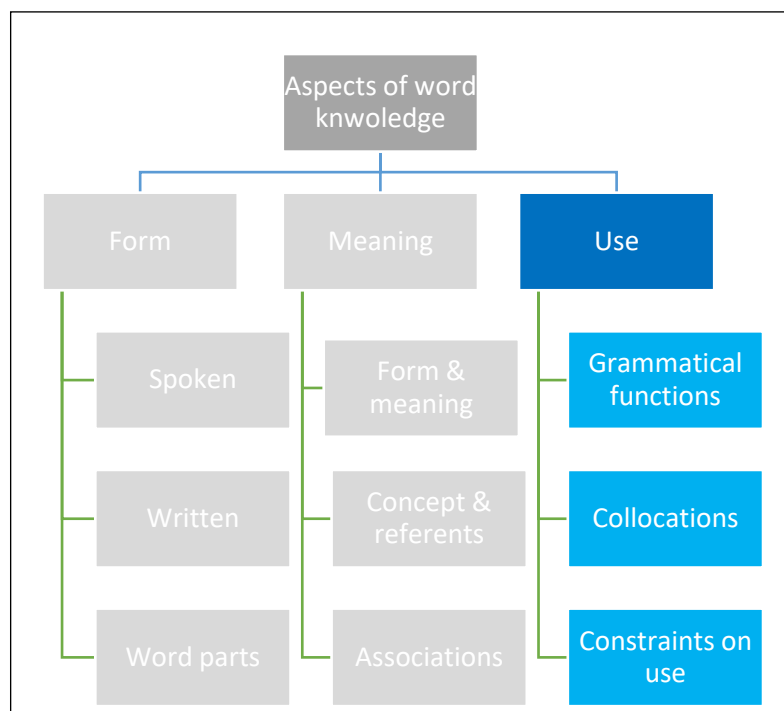


Figure 7: Use (Nation 2013: 48)

This already leads to the last category, namely knowledge about the usage of a word. Nation (2013: 82) states that “in order to use a word it is necessary to know what part of speech it is and what grammatical patterns it can fit into”. This means that grammatical aspects, such as tense, irregular forms or word order need to be investigated. Moreover, learners should be aware that words tend to co-occur with other words, forming collocations (Nation 2013: 83). Durrant and Schmitt (2010: 163) acknowledge the importance of collocations for language fluency. Studies (Siyanova & Schmitt 2008; Durrant & Schmitt 2010) have suggested that learners are capable to produce a large number of correct collocations in their L2. However, Siyanova and Schmitt (2008: 429) discovered that L2 speakers’s collocation use differs from native speakers’ use in intuitive use of collocations and speed of processing, since native speakers outperformed L2

learners. Furthermore, Nation (2013: 83) states that collocations are only one part of knowing how to use a word. He argues that learners need to build up context knowledge, which can be grouped in three subdivisions, namely situational, topical and local context (Nation 2013: 83). While the local context is created by the surrounding lexical items of a word, topical and situational context have a broader focus (Nation 2013: 83). Topical context is concerned with the content that speakers are negotiating and situational context is concerned with the surroundings of the communication, such as location, speaker relationship or type of interaction (Nation 2013: 83). This leads directly over to the last point that learners need to consider, namely sociolinguistic constraints of word usage (Nation 2013: 84). For example, the formality of a situation or the relationship of two speakers can influence and restrict word choice. Moreover, certain words can be culturally biased, having different semantic connotations or even being impolite. For instance, in Thai *pig*, *fatty* or *shrimp* are used as nicknames (Nation 2013: 84). However, these would not be considered to be appropriate by speakers of other cultures.

All the aspects of word knowledge mentioned above have an influence on the learning burden of a word. The difficulty of a specific word differs, however, from learner to learner (Nation 2013: 44). These differences in perceived learning burden are caused by the divergent prerequisites of individual learners, such as their existing background knowledge or their knowledge of other languages (Nation 1990: 33). Especially the L1 has been proven to have a considerable effect on the learning burden of words (Nation 1990: 33). For the spoken form of a word the learning burden will be relatively low if the words “use the same sounds and arrangement of sounds as in the learner’s mother tongue” (Nation 1990: 35). The same accounts for the written form of a language. If, for instance, a Chinese person would try to learn English, the orthographical form will be difficult to grasp because the learner is not used to the Roman script. In general, the more predictable the language to be learned is for the student the weaker the learning burden will be (Nation 1990: 35).

2.4.2. Learning styles and learning strategies

Recent years of investigation have shown an increasing interest in both learning styles and learning strategies as well as in their impact on language proficiency or learning success (Griffiths 2008: 49). Nevertheless, both terms remain rather elusive, since researchers cannot agree on a suitable definition and use various divergent terms as

synonyms (Griffiths 2008: 83). Amongst other scholars Curry (1990: 53) particularly points at the fuzzy boundaries between the concepts of style and strategy. Since this study is mostly concerned with learning strategies, it appears to be relevant to clarify the relationship in which learning styles and learning strategies stand. Although various successful models for learning styles exist (Dunn & Dunn 2005; Reid 1987) this study is going to apply Curry's (1983) model, which is not the most recent but the most suitable for its purpose. Curry's (1983: 7) metaphor for learning style is an onion with three distinctive layers:

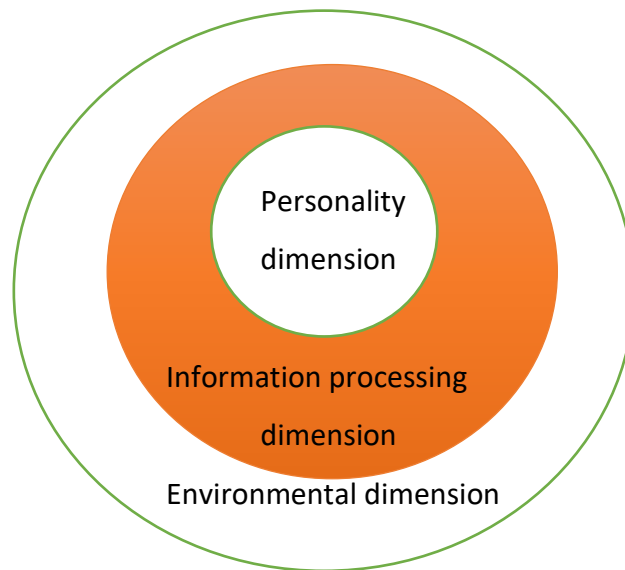


Figure 8: Curry's model for learning styles (1983: 8)

Similar to the skin of an onion, which is in contact with the surrounding earth, the outer layer of the model is in direct contact with the learning environment (Curry 1983: 8). For instance, it is concerned with the spatial conditions of learning or the availability of learning material. The second layer considers all the learner's preferences concerning information processing (Curry 1983: 8). Therefore, this dimension pays attention to the learning behaviour of students and can be equated with learning strategies. The inner core of the metaphorical onion deals with learner's personality traits and is thought to be relatively stable (Curry 1983: 8). According to this model it can be argued that learning strategies are part of the learning style of a person. To take it one step further, the functions that the inner layer and the outer layer have for the second layer might be explored. Depending on how fruitfully personality (inner layer) and environment (outer layer) collaborate, learning strategies might be expedient. However, the second layer might also be used to take countermeasures against inhibiting environmental conditions

or personal traits. Therefore, learning strategies will be seen as part of an interactive model of learning style.

2.4.2.1. Language learning strategies

Schmitt (1997: 199), Griffiths (2008: 83) and Tacac (2008: 47) all criticise the non-transparent use of the term language learning strategies in existing literature. Therefore, a closer investigation of different approaches is necessary to formulate a suitable working definition for this paper. Nation and Gu (2007: 82) define learning strategies as “a series of actions a learner takes to help complete a learning task”. This relatively broad definition gives rise to the question as to what is subsumed under the term learning task. The answer can be found in Rubin (1987: 19) who states that learning strategies are “what learners do to learn and do to regulate their learning”. Therefore, not only the effect of knowledge enhancement but also the processes involved in regulating learning are affected. Ellis and Beaton (1995: 148) further distinguish between mental and behavioural processes in learning, which means that strategies might be purely cognitive or may require an action. Nevertheless, it remains questionable whether a strategy can be solely mental or behavioural. On the contrary, strategies appear to be mostly cognitive and behavioural, since hopefully an action is preceded by thoughts. Schmitt (1997: 203) expands on Rubin’s (1987) and Ellis’s and Beaton’s (1995) approaches by claiming that a learning strategy is “the process by which information is obtained, stored, retrieved and used” and therefore appears to validate Ellis’s and Beaton’s focus on behavioural strategies the definition.

Wenden (1987: 6-7) introduced another notion to the debate, namely strategic knowledge. He states that learning strategies “[refer] to what learners know about the strategies they use”. This implies that learning strategies need to be chosen consciously. Oxford (1990: 1) appears to validate this thought, since she states that “learning strategies are tools for active and self-directed involvement”. However, she adds that learning strategies can become unconscious after choosing them with a certain amount of practice. This stance is shared by Cohen (1998: 19) and Giffiths (2008: 85) but is refuted by Ridley (1997: 30) and Purpura (1999: 24), who claim that strategies might be unconscious as well. The controversy is triggered by the lack of specification of the term conscious by both Griffiths (2008: 85) and Cohen (1998: 19). If they subsume direct and indirect learning under this term, their argumentation is plausible. For instance, if a learner uses

extensive reading for vocabulary training, which may be classified as indirect learning, the degree of awareness concerning the beneficial nature of extensive reading for vocabulary is decisive. If the student consciously chooses extensive reading to increase vocabulary, this might be classified as a learning strategy. However, if the student just incidentally happened to improve his/her vocabulary because he/she likes to read, extensive reading would not be counted as a vocabulary learning strategy in my study. As a consequence, Wenden (1987), Cohen (1998) and Griffith (2008) raise an important point, namely that strategic knowledge is required to be able to make conscious decisions, which are then referred to as strategies.

Lastly, Oxford (1990: 8) states that language learning strategies adhere to certain features: Firstly, all language learning strategies are problem oriented (Oxford 1990: 11). They work towards the primary goal of communicative competence (Oxford 1990: 8). Secondly, they alter the roles of learner and teacher. While the learner is getting more independent, the teacher needs to take on different social roles, such as helper and advisor (Oxford 1990: 10). The degree of guidance that a student needs might also depend on factors such as learning style and personal traits as motivation. Oxford (1990: 11) calls this “action bases” of a strategy. Students who display a lack of strategic competence due to such reasons can however be taught how to adapt strategies effectively to compensate (Oxford 1990: 12). This highlights the flexible nature of learning strategies (Oxford 1990: 13). As mentioned above, Oxford (1990: 12) believes that learning strategies are mostly consciously but can become unconscious through a certain amount of practice. The choice of a strategy might be influenced by a variety of factors, such as gender, age or degree of awareness (Oxford 1990: 13) However, it is important to note that also conscious learning strategies might not be visible to an observer, since they can be mental processes (Oxford 1990: 12). Moreover, learning strategies might influence learning directly and indirectly (Oxford 1990: 11-12). This point will be discussed in more detail in chapters 2.4.2.2 and 2.4.2.2.4.

Drawing these observations together it can be stated that learning strategies are consciously chosen, teachable, flexible, cognitive and/or observable behavioural processes to regulate the planning, obtaining, storing and practicing of information to ensure reaching the goal of communicative competence.

Having formulated a working definition of learning strategies for this study, the features mentioned can be transferred to one particular group of learning strategies, namely vocabulary learning strategies. Gu (2003: 73) states that a vocabulary learning strategy includes thoughts about the task itself, plus the planning, use, monitoring and evaluation of learning behaviours to learn new vocabulary. Moreover, Gu and Nation (2007: 82) indirectly take up Curry's model of learning style as can be seen in Figure 9 by stating that learner, task, context and strategy are linked. Nevertheless, they condense the three layers of Curry's model to two layers with an interactive nature.

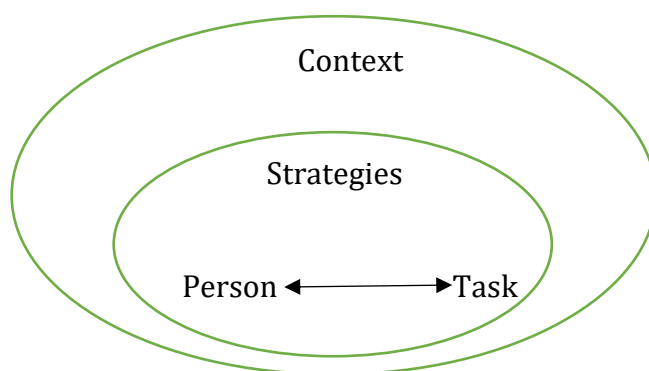


Figure 9: Gu's and Nation's model (2007: 82)

They extend previous thoughts about the interrelation of the component parts by noting that some vocabulary items are more oriented on the learner's personal traits, while others focus on the context or the task itself. Therefore, they consider layer 1 and 3 of Curry's model and they add to layer 2, which is concerned with learning strategies, by showing that the learning task itself is also an influential factor. Blending these two approaches into one picture, a physical metaphor might be used for a new approach of a model.

In order to highlight the interactive nature, the relation between the components of the model could be compared to the manual task of forging. The people using the hammer have different strengths, experiences and talents. The task can be of varying difficulty depending on the material the person is trying to form. If it is iron it will be relatively easy but if it is steel it will be a more challenging task. The context is the physical rules and realities of earth, such as gravity, which influence success as well. The hammer chosen is the strategy. Depending on the strength of the person using it, the length of the shaft might vary. If the person is relatively weak a longer shaft can be employed in order to take advantage of the leverage.

In the following a taxonomy of vocabulary learning strategies will be developed, which is a compilation of strategies found in primary and secondary literature.

2.4.2.2. Taxonomy of vocabulary learning strategies

Reviewing the existing language learning strategy taxonomies, one can see that no consensus has been reached (Schmitt 1997: 203), which means that studies are difficult to compare, since they are based on different vocabulary learning taxonomies. To contribute to a solution of the controversies in the field this paper combines existing vocabulary learning taxonomies. Thereby it relies heavily on Schmitt's (1997: 207-208) taxonomy developed on the basis of his study and on the basis of Oxford's (1990: 15) division of strategies. Oxford (1990: 15) distinguishes between six strategy groups, namely Memory strategies, Social strategies, Affective strategies, Metacognitive strategies, Compensation strategies and Cognitive strategies. These are again grouped into two superordinate groups termed direct (Compensation, Cognitive and Memory strategies) and indirect strategies (Social, Affective and Metacognitive strategies). Direct strategies are all strategies that are used for "dealing with new language", while indirect strategies are strategies for "general management of learning" (Oxford 1990: 14-15). However, this division will not be applied in this strategy, although the terms direct and indirect strategies are used but with another meaning as explained in chapter 2.4.2.2.4. In his vocabulary taxonomy Schmitt (1997: 205) claims to borrow three of the strategy concepts of Oxford, namely Cognitive, Social and Memory strategies. While the distinction between social and memory strategies will be adopted for this taxonomy the concept of Cognitive strategies was not used as will be justified later.

So far no research has been done on this new combined taxonomy, which would be desirable in future investigations. Figure 10 illustrates the structure of the taxonomy used in this study.

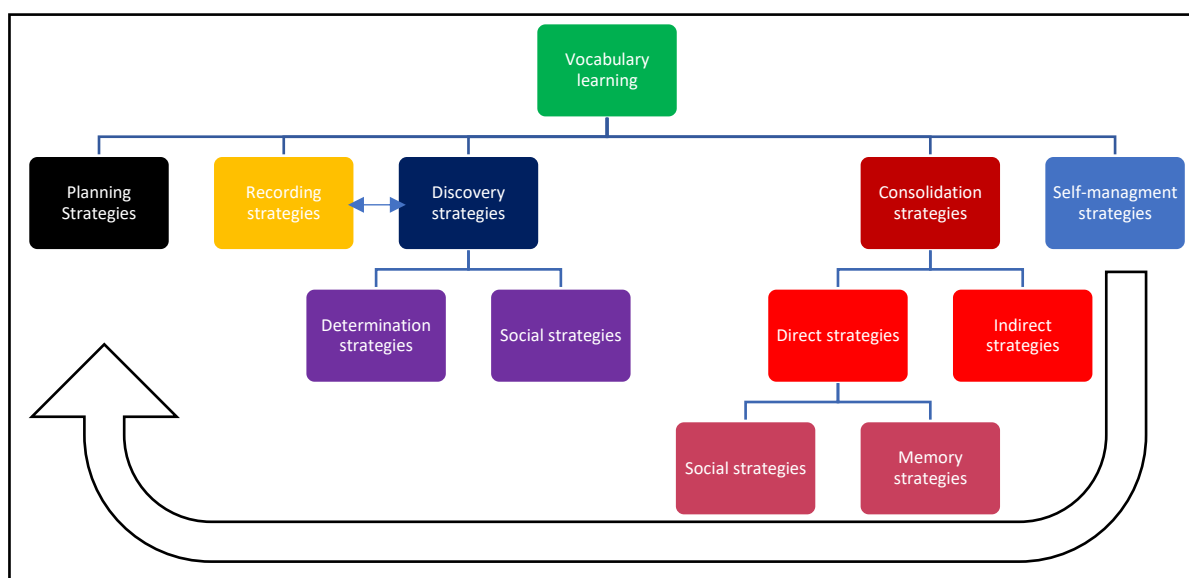


Figure 10: Taxonomy of vocabulary learning strategies used in this study

2.4.2.2.1. Planning strategies

Following Nation’s (2013: 329) example, the first vocabulary learning strategies for independent learners are considered with planning vocabulary learning. This is based on the assumption that the taxonomy is mostly designed for active learners of English who wish to broaden their vocabulary knowledge. Therefore, this first category is concerned with processes that should ensure successful learning, such as planning which words, aspects of words and learning strategies should be considered (Nation 2013: 329). This category could be compared to Oxford’s (1990: 17) Meta-cognitive strategies, since they involve arranging and planning strategies for learning as well. However, other strategies subsumed under this category such as “evaluating learning” do not fit in the concept of planning strategies in this taxonomy but will be grouped in a different category (Oxford 1990: 17).

According to Paul Nation (2013: 328) planning vocabulary learning consists of four separate steps. Firstly, learners need to make a well-considered choice concerning the vocabulary items that they want to learn. Secondly, students reflect on which aspects of word knowledge they need to concentrate on. Thirdly, suitable vocabulary learning strategies need to be found and lastly, vocabulary repetition and the time necessary to acquire the items need to be planned.

Gu and Johnson (1996: 659) highlight the fact that strategic decisions concerning the words to be studied are a crucial factor for success. They (1996: 659) justify their

statement by providing convincing evidence that students' selective attention to items contributes to their progression towards a learning goal. This can be complemented by the Involvement Load Hypothesis, which claims that the higher the involvement of students is in a task, the more effective learning is going to be (Laufer & Hulstijn 2001: 21). However, this raises the question as to how students can formulate successful learning goals which are triggered by their needs. One possibility would be to employ Nation's Vocabulary Levels Test, which "can be used to measure whether the high-frequency words have been learned, and where the learner is in the learning of academic and low-frequency vocabulary". Therefore, it indicates which type of words the student should focus on (Nation 2013: 36). This and various other similar tools based on research, such as the Words and Phrase Tool by Townsend and Kiernan (2015: 116), which helps independent learners to discover their weaknesses in vocabulary knowledge, are available for free on the internet. Moreover, focusing on learning already existing lists which suit the aims of a course or personal aims is an option too. For instance, medical professionals could focus on discipline-specific word lists, which offer a compilation of technical vocabulary. In addition, learning material, such as practice books mostly offer a compilation of vocabulary items in the form of a list. However, one should check whether the vocabulary list could be improved before using it.

What is most important when choosing vocabulary items to study is that learners use a strategy for choosing a word and can explain why they learned a particular word. It is crucial that students are familiarised with these possibilities before starting to work independently (Nation 2013: 329). Otherwise the effort will probably not lead to the success in vocabulary learning expected.

Chapter 2.4.1. of this paper is focused on the aspects of word knowledge which should or might be considered in order to learn vocabulary items successfully. In order to learn goal-orientedly and effectively learners should be aware of these features and should be proficient enough to make a suitable choice as to which aspects to focus on to reach their aims (Nation 2013: 329). For instance, a student aiming to improve their academic writing, such as the participants in this study do, should focus on the written form of words and not primarily on pronunciation.

A response often heard to teachers' offers or insistence to try out a new vocabulary learning method is: Why should I use a new way of learning vocabulary if I already have

an efficient method? This question is actually easily answered. Sanaoui (1995: 26) and Gu and Johnson (1996: 668), who investigated 50 students attending an ESL vocabulary learning course in Canada, report that the number of strategies used by learners correlates positively with their language proficiency. A more recent study by Hong-Nam and Leavell (2006: 402) investigated 55 students of a language learning institute for pre-admissions university ESL students with mixed cultural background grouped into three language ability levels, namely beginners, intermediate and advanced according to the level of classes they were taking. They found that intermediate learners tended to use more strategies than the other two groups. Nevertheless, the more strategies students used, the faster they improved (Hong-Nam & Leavell 2006: 410). These results were partially contradicted by Ansarin, Zohrabi and Zeynali (2012: 1842), who investigated the relationship between language learning strategies and vocabulary size amongst Iranian EFL learners. They state that participants with advanced language abilities (TOEFL level) used more strategies and had a larger vocabulary size than lower ability learners (Ansarin, Zohrabi & Zeynali 2012: 1847). Therefore, Ansarin, Zohrabi and Zeynali's study partially negates Hong-Nam's and Leavell's research by stating that advanced learners used the most strategies. However, Ansarin's, Zohrabi's and Zeynali's (2012: 1842) study appears to be more reliable because in contrast to Hong-Nam and Leavell they used an acknowledged placement test from Oxford and Cambridge University from 2001. A direct positive correlation between the amount of strategies and learning progress is highly probable, as John Barcroft (2009: 74) reported in his own study. However, divergent results could also be explained by the different contexts of the studies, since all studies were conducted in different cultures and only Leavell and Hon-Nam's study had participants at beginners University level, which therefore would be most comparable to the setting of this investigation.

Griffiths (2008: 89) amongst other scholars already insists on treating such results with caution. Although she (2008: 89-90) also reported a positive correlation between vocabulary learning strategies and language proficiency, she states that other studies such as Vann and Abraham's (1990: 177) investigation showed that also weak students might use a great variety of strategies but very unsuccessfully. This ties in with what Nation (2013: 329) reports, namely that "successful strategy users need a strategy for controlling their strategy use". Successful vocabulary learners should therefore not only have a large

repertoire of vocabulary learning strategies but also need to be able to choose, adapt and combine them efficiently in a given situation.

To learn vocabulary permanently, revision after the first noticing of a word is necessary. Learners should therefore know how often they need to retrieve a word in order to remember it permanently. The more often a word is met receptively or is used productively, the better the memorization is going to be. Webb (2007: 62) investigates how repetition and word knowledge are linked amongst 121 EFL students in Japan. Learners were divided into four groups. All groups got different pages to read in which the target word was used. However, group one only encountered the target word once, group two three times, group three seven times and group four ten times. After all participants had read their assigned text a vocabulary test was completed. He found that participants were able to remember more aspects of a word, such as formality or collocations, the more often they revised it. However, he stated that after ten encounters that his subjects had with the vocabulary item knowledge was still not completed (Webb 2007: 62). To plan revision Nation (2013: 329) recommends the use of “increasingly spaced retrieval”. He argues that learners can use an informal system of reusing vocabulary material or that they could use a review system like computer programs (Nation 2013: 329).

2.4.2.2.2. Recording strategies

After a hopefully successful planning stage, students move on to store the knowledge physically with the help of specific strategies. For instance, the production of word lists, word cards or vocabulary notebooks would be grouped in this category. Schmitt (1997: 206) equates the strategies subsumed under this category with Oxford’s (1990: 15) cognitive strategies. He argues that like Oxford’s primary definition of cognitive strategies, the methods used at this stage of vocabulary learning involve a manipulation of the language by the learner (Schmitt 1997: 205). Although his argumentation appears to be valid on the surface his statement does not get to the heart of Oxford’s idea of cognitive strategies, since they involve not only manipulation of the language but “practicing, receiving and sending a message, analysing and reasoning” (Oxford 1990: 17). Therefore, this taxonomy will not use this terminology but has termed this group of strategies recording strategies.

Similar to Nation (2013: 331) Schmitt (1997: 215) mentions unstructured written recording of a word, later explained in more detail as part of rote-learning, as well as word lists, flash cards or vocabulary notebooks as the most prominent recording strategies. Moreover, Schmitt (1997: 216) extends Nation's list by mentioning the tactile method of sticking sheets with words on them onto objects. However, learners might even decide against recording vocabulary items at all and just revise them orally. In the following most of these recording strategies will be explained in more detail and will be checked for their effectiveness. It is important to mention that separating the benefits of recording strategies from the advantages for vocabulary learning is difficult since recording strategies aim at achieving the optimal vocabulary learning success. This means that recording strategies are often also judged on the basis of the amount of vocabulary items learned by using them and the time required to learn vocabulary with the help of this recording strategy.

Nation (1990: 126) states that although using word lists has become unpopular amongst teachers over the last few years, it still remains a fashionable strategy for independent learners. All types of word lists have the benefit of being able to record and learn a large amount of words in a relatively short time. On average, learners master 30 words per hour using a list with translations, while high performance students even pick up 100 words per hour. These words can still be retrieved after several weeks (Nation 1990: 126). Nevertheless, some types of lists enhance learning more than others. Nation claims that a combination of the new foreign word with a L1 cognate is more beneficial in a list than the use of L2 synonyms or definitions for word lists. Oxford and Crookall (1990: 10) distinguish between unpaired lists and paired lists of which the former offers only the L2 words without further information and paired lists, which conform to Nation's preferred listing method of using L1 equivalents.

However, Nation (2013: 437) states that he prefers flash cards over word lists. He defines learning from word cards as "the formation of associations between a foreign language word form and its meaning" (Nation 2013: 437). Learners hereby produce cards with unknown words on them (Nation 2013: 437). Optionally they can provide additional information on the cards, such as L1 translations, definitions or synonyms. All types of flash cards have three clear advantages over word lists according to Nation and Gu (Nation & Gu 1990: 88; Nation 2013: 437). Firstly, learners can recall the meaning of a word without the potential danger of seeing the answers. Secondly the order of the words

can be changed, while the word list is static. This has the benefit that learners cannot remember the words by the order in which they occur so easily instead of remembering their meaning, as it is the case with word lists. Thirdly, learners may group word cards according to their difficulty and can adapt revision accordingly. Regarding the first point one needs to mention that this flaw of word lists could be easily removed if students cover the column with the answers with a sheet or fold the word list in the middle to have the answers on the back. Concerning the third point Nakata (2008: 7) criticizes that some learners might not be self-reflected enough to group and revise words effectively. As a solution, she promotes flash card computer programs, which undertake this task for the learners (Nakata 2008: 7). Nevertheless, it remains questionable whether this evasion is a permanent solution for this problem or whether computer programs rather enforce and promote the principle of avoidance. One could compare the situation to a simpler task. For instance, if a baby is learning to walk but it stumbles sometimes before it reaches perfection nobody will consider giving the baby a wheelchair to protect it from falling and to help him to learn, since the baby will never learn how to walk then but only how to move with the wheelchair. The same accounts for learning how to organise and learn with word cards. It is normal that students experience difficulties or even fail sometimes but if we give them a tool that does the task for them they will never learn how to do it. Therefore, the application of computer programs should always be well considered and based on research. One positive example is the implication of multimedia flash cards by Aronin and Haynes-Smith (2013: 33). Based on precise research on effective instructional strategies, the relationship between vocabulary and reading comprehension, and the effect of imagery in supporting vocabulary recall they developed PowerPoint flash cards which were employed to increase the technical vocabulary knowledge of students in certain fields (Aronin & Haynes-Smith 2013: 34-35). Therefore, this project would be a positive example for the application of multimedia tools for vocabulary learning.

Having named the benefits of flash cards, it should be admitted that there have been voices of criticism as well. Oxford and Crookall (1990: 9) have subsumed flash cards under the construct of decontextualized learning, which has been heavily criticised. According to Oxford and Crookall (1990: 9-10) decontextualized learning can be described as “techniques [...] that remove the word as completely as possible from any communicative context that might help the learner remember and that might provide some notion as to how the word is actually used as a part of language”. According to this definition other

strategies fall under this concept as well, such as learning from word lists or using certain types of reference sources. Nation (2013: 438) names four main points of critique against flash cards:

Learning from word cards is not good for remembering.

Learning from word cards does not help with the use of the word.

Learning from word cards only provides explicit knowledge which is not the kind of knowledge needed for fluent use.

Deliberate learning can only deal with a small number of the words which need to be learned

Concerning the quality of remembering words, which were learned in a decontextualized context Nation mentions several studies (Bahrick 1984; Bahrick & Phelps 1987; Thorndike 1908; Beaton et al. 1995) which prove that vocabulary items were available a long time after the first contact. However, all the studies named are relatively old and should be replicated with recent data to be validated.

Regarding point two Nation (2013: 439) counters that teachers often appear to “underestimate learners’ capacity for the initial learning of foreign vocabulary”. Depending on the size of their working memory, students were able to acquire between 9 and 58 word pairs per hour (Thorndike 1908: 127). A study by Unaldi et al. (2013: 91) from different universities in Turkey even showed that decontextualized learning can have more effect on vocabulary recognition than contextualized learning. They investigated 69 participants who were divided into three groups. The same vocabulary items were taught to all subjects but group one received contextualized instruction, while group two was exposed to decontextualized instruction. The third group tried a corpus-informed approach. The results indicated that the most progress in vocabulary recognition was made by the decontextualized learners, followed by the corpus-informed instruction group and lastly the contextualized group. These findings were validated and extended in a study by Choi, Kim and Ryu (2014: 228), who investigated Korean high school students. In the eight week experiment the participants had to acquire 30 English words incidentally from reading stories and another 30 words by learning from a decontextualized word list. After five weeks, implicit as well as explicit lexical knowledge was measured. The results showed that decontextualized learning was more sufficient for establishing explicit lexical knowledge (Choi, Kim & Ryu 2014: 230). These findings

provide a nice transition and proof for Nation's (2013: 441) reaction to Oxford and Crookall's (1990: 12) critique that learning from flash cards is not sufficient for learners to be able to use the words. He argues that flash cards provide learners with explicit knowledge of the basic concept of a word, while learning from context adds several shades of meanings to this basic concept (Nation 2013: 441). Therefore, he proposes that learning from word cards should be viewed as a complementation of learning from context. The explanation above and results additionally give a sufficient answer to criticism number three. Since decontextualized learning enhances explicit as well as implicit knowledge (Nation 2013: 443), learning from flash cards also facilitates the development of fluency. Moreover, Biemiller and Boote (2006: 51) falsify the fourth point of critique. They argue that decontextualized learning can result to an increase of approximately 400 words a year if students learn 10 to 20 words a week.

Lastly two other recording strategies can be mentioned, namely the labelling of objects with the help of post-its and the making of a vocabulary notebook. While the latter exists in unlimited variations and is suitable to collect all aspects of word knowledge, the former might be rather limited since objects are only nouns. Therefore, this method might be more suitable for beginners.

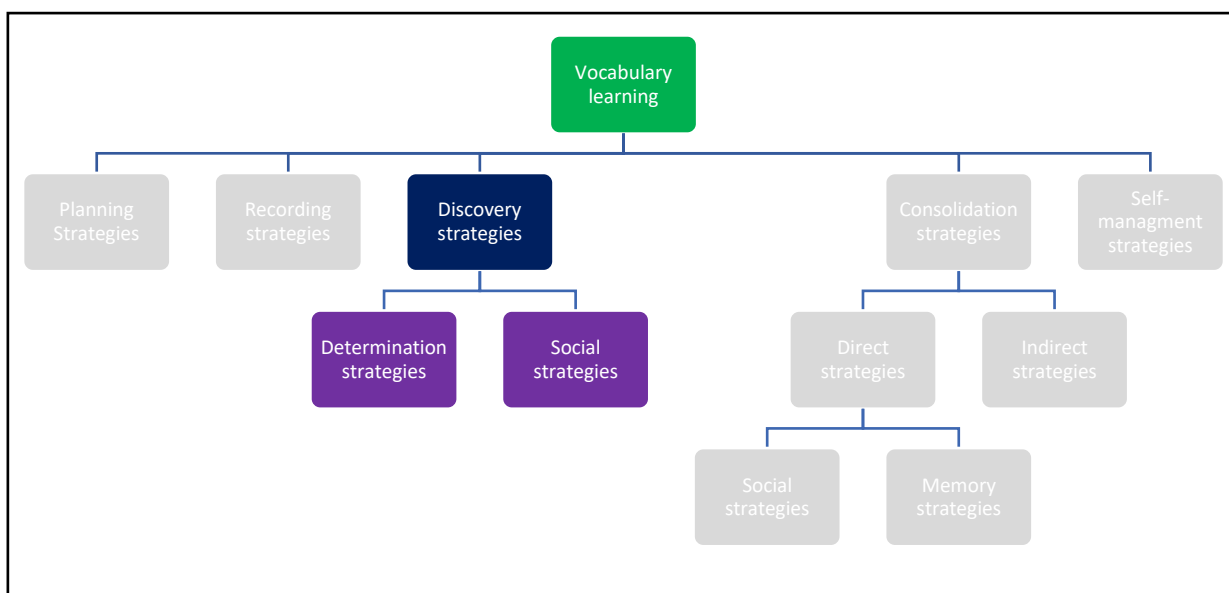


Figure 11: Discovery strategies

2.4.2.2.3. Discovery strategies

After students found a successful recording system for themselves they might encounter more new words, which they might want to add to their previous vocabulary knowledge.

For these first encounters with an unknown word, discovery strategies are applied. Nation (1990) has defined discovery strategies as methods to find out what a word means for the first time. Therefore, this category includes strategies such as consulting a reference source or analysing the affixes and roots of a word. Schmitt (1997: 206) adds that discovery strategies can be divided into two sub-categories, namely determination and social discovery strategies. While determination strategies are used by the learner without any help from other people, social discovery strategies involve an attempt to discover the meaning of a word through interaction with other people (Schmitt 1997: 206). For instance, consulting a reference source would be a determination strategy, while asking a classmate for the meaning of a word is a social strategy (Schmitt 1997: 206). This conforms to Oxford's (1990: 17) definition of social strategies, who subsumes asking, cooperating and emphasising under this strategy group. So far it has been assumed that the target audience of the taxonomy is the active language learner. However, all language users will have to apply discovery strategies, since they might read the local newspaper of a political party for the upcoming elections or might simply browse through a lifestyle magazine and encounter a word that they do not know. Consequently, they will apply discovery strategies at hand to understand the meaning of the word or they might ignore it. What differentiates the language users from the active language learners is that they have not spent any effort on planning or recording and that they will not attempt to learn the vocabulary as described in the next step, as the active learner will do.

Discovery strategies are divided into social strategies and determination strategies. Social discovery strategies are easily named and explained. They comprise negotiation about the meaning of a word with another person with the same or different proficiency level, such as a classmate, or a lecturer and/or native speaker. Determination strategies on the contrary need more explanation. Nation (2008: 98) mentions four determination discovery strategies, namely comparison with L1 cognates, analysis of available pictures, dividing a word into its parts and analyse them and guessing from context. Schmitt's (1997: 208) taxonomy can be used to supplement this list with analysing the grammatical context of the word and the use of diverse reference sources. As an addition, checking for defining words or phrases in the context and the skipping of an unknown word can be named. It needs to be mentioned that not all discovery strategies will be described in the same detail in the following.

The analysis of the grammatical surroundings of the word, word parts, checking for defining words or phrases, analysis of available pictures and ignoring of an unknown word are quickly explained. An unknown word is mostly embedded in a sentence, which means that maybe the meaning of the word becomes accessible by looking at the semantic content of the immediate context of the word. Moreover, a grammatical analysis of the sentence in which the unknown word occurs might reveal which grammatical category the word belongs to. This can facilitate the process of guessing the meaning of the word (Schmitt 1997: 208). The same accounts for word parts. Schmitt (1997: 208) states that learners might be able to infer the meaning of an unknown word by taking a look at its roots and affixes. Another method would be to extend the amount of context paid attention to by searching the text for definitions of the word. A similar but more obvious approach would be to see whether the text contains any pictures that could provide information about the meaning of the word. However, learners might also decide that they will simply ignore the word and keep on reading.

A very commonly used discovery strategy is the procedure of searching for cognates in the L1. Schmitt (1997: 209) defines cognates as “words in different languages which have descended from a common parent word, such as *Mutter* in German and *mother* in English”. These connections between known languages and languages to be learned can but must not be useful for successful guessing and remembrance of meaning (Schmitt 1997: 209).

This leads directly to the next strategy, namely guessing from context. Guessing the meaning from context is a strategy which can be applied by learners with at least 3000 words vocabulary size (Nation 1990: 160). By interpreting the text surrounding an unknown word its meaning may be inferred. However, it is crucial that learners know at least 95 percent of the total running words of a text to have a chance of guessing correctly (Nation 2013: 352). Moreover, students differ in their given abilities to infer meaning (Nation 1990: 160). Therefore, guessing meaning from context is not just an inborn ability but a strategy which can and needs to receive specific training (Nation 2008: 101). However, it is important to acknowledge the limitations of guessing from context. Nation (2008: 77) concludes that “the bits of information gathered about each word may be small, and because the reader’s main attention is on the story, this knowledge will not be strongly established”. Therefore, it is important to use other sources of information in addition to discover the meaning of a word.

A common reference source is the dictionary. Several scholars (Schmitt 1997; Bruton 2007; Nation 2013) have confirmed that dictionary use is beneficial for vocabulary learning. Nation (2013: 372) emphasizes that the choice of a dictionary is an important decision, since the quality of the reference source influences the quality of learning. While Schmitt (1997: 209-210) states that the bilingual dictionary is still students' preferred choice over the monolingual dictionary Nation (2013: 424) states that monolingual dictionaries normally provide more information than a bilingual dictionary. Moreover, bilingual dictionaries are often criticized for their enforcement of translation and the idea of a one-to-one relationship of languages. It is also claimed that they give little information on the use of a word (Nation 2013: 424). However, Nation (2013: 425) argues that also a bilingual dictionary can have its advantages since it is presenting information in an accessible manner and can be used bi-directionally. A study by Atkins and Varantola (1997: 19) even proved that learners had been more successful in finding the correct meaning of a word with bilingual rather than with monolingual dictionaries. As a solution, Nation (2008: 114) brings forward a convincing distinction. He states that to use a monolingual dictionary learners already need a vocabulary size of 2000 words, since this is the "controlled defining vocabulary" of the dictionary (Nation 2008: 114). Therefore, he suggests that beginners should rather use a bilingual dictionary, while more advanced learners can make use of monolingual dictionaries.

In recent years online dictionaries have become more prominent. However, feelings about this innovation are mixed. While some scholars (Li 2010) praise the benefits of online dictionaries, others remain more sceptical. Atipat Boonmoh (2012: 43) concluded, after discussing the results of his study, that online reference tools have advantages but also disadvantages. Again Nation (2008: 114) solves the controversy by stating that also in this case the quality of learning is dependent on the quality of the reference source. Additional possibilities of reference sources would be to utilize existing word lists or flash cards on the market or to make use of language corpora, which are a valuable resource of information on natural language use.

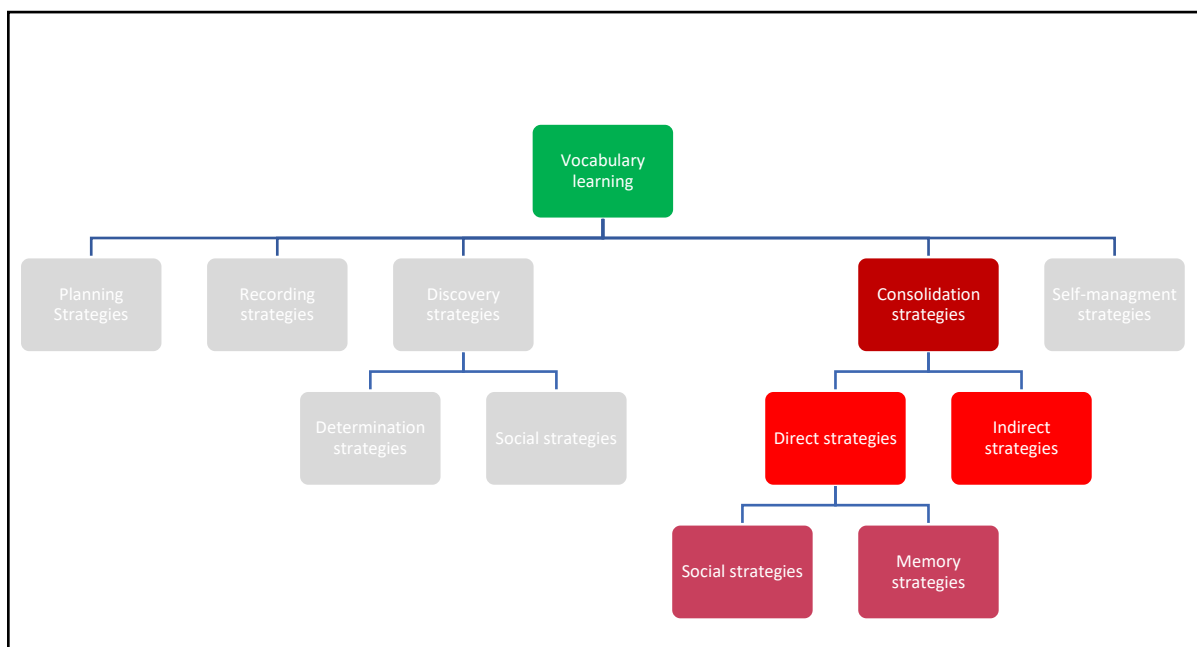


Figure 12: Consolidation strategies

2.4.2.2.4. Consolidation strategies

Strategies to actively learn vocabulary are named consolidation strategies and were again inspired by Paul Nation (1990). Consolidation strategies broadly include all strategies available for the active effort of learning and remembering a word, such as connecting a word to a personal experience, synonyms or visual representations to remember it (Schmitt 1997: 211). However, research has shown that learning vocabulary does not have to occur by active manipulation of or focussed attention to words but can also be incorporated in other tasks (Nation 2013: 348). This distinction is termed direct and indirect learning in the following. While direct learning involves a conscious focus on a word, indirect learning includes strategies such as learning vocabulary through reading or interaction (Nation 2013: 302). This distinction therefore differs significantly from Oxford's (1990: 16) interpretation of direct and indirect learning, who sees indirect language learning as detached from language material and focused on overarching learning functions, such as planning strategies. Nevertheless, in this study indirect consolidation strategies were only counted as such if learners reported that they consciously decided to use a skill, such as reading or listening, with the goal of enhancing their vocabulary knowledge. In contrast, to the indirect consolidation strategies, direct consolidation strategies can be further divided into social and memory strategies (Schmitt 1997: 206). Both involve direct attention to a word but social consolidation strategies are applied in pairs or groups, while memory strategies are used alone. For example, a social

consolidation strategy would be to study in a group or to act out role plays to remember words (Schmitt 1997: 207).

Direct social strategies include studying in a group or with a native speaker (Schmitt 1997: 207). These offer all the benefits of learning in groups. Schmitt (1997: 211) summarizes and updates the advantages collected by another researcher (Dansereau 1988). He states that learning in groups

promotes active processing of information and cross modelling/imitation; the social context enhances motivation of the participants; cooperative learning can prepare the participants for 'team activities' outside the classroom; and because there is less instructor intervention, students have more time to actually use and manipulate language in class.

A more complex direct social strategy which can be useful for learning a word is role play. The procedure of acting specific roles can be utilized for vocabulary learning, since "role play potentially offers an enjoyable way to encourage students to use vocabulary appropriately in simulated authentic situations" (Alabsi 2016: 227). To prove the effectiveness of this strategy Alabsi (2016: 230) conducted a study with 40 female secondary school students. She asked one group to study vocabulary through role play and the second group to use conventional methods. The post-test showed that students using the role play strategy had a significantly higher score than the control group (Alabsi 2016: 232). Therefore, the role play strategy appears to be a successful choice at least for a suitable group of learners. It can be assumed that the same principle can be found in the preparation of talks incorporating new vocabulary.

Memory strategies are one of the most researched groups of vocabulary learning strategies (Gu & Nation 2007: 90). This might explain why Schmitt (1997: 207) lists an incredible number of different memory strategies used for consolidation in his taxonomy. Since some strategies yielded more interesting results than others or were simply more prominent in research only a selection will be discussed in more extent in the following.

A number of memory strategies works with different visualising strategies. The simplest strategy is to group words spatially in a list, in scales or feature grids. Words are often memorized with the help of rote-learning, which includes revising the words in the list regularly by writing them down or by speaking them out loud. However, the meaning of words can also only be visualised cognitively, such as when picturing the meaning of a

word or the graphical form of the word. Several scholars (Atkinson 1972; Laufer & Hulstijn 2001) heavily criticize the unorganised method of simple rote-learning. However other findings suggest that rote-learning can have a positive effect as well. Omid and Rashidi (2011: 146) argue on the basis of their study that certain types of learners have managed to utilize this simple strategy with astonishingly positive results. This is supported by Schmitt (1997: 215), who remarks that rote learners can reach high proficiency levels. A study by Gu and Johnson (1996) provides more detailed results. They claim that oral repetition had a low positive correlation with language proficiency and vocabulary size, while visual repetition was even a strong negative predictor (Gu & Johnson 1996: 655). Yang and Dai (2011) and Tan (2011) all try to explain these controversial results. They state that rote learning is strongly related to cultural background (Yang & Dai 2011: 62; Tan 2011: 124). While Asian cultures are mostly focused on rote learning and manage to apply this strategy successfully, western cultures tend to dislike rote learning and to use it insufficiently (Tan 2011: 125). Therefore, the degree of proficiency reached with the help of rote learning might be dependent on the cultural embeddedness of the strategy in the culture of the learner. Relying on these results, rote learning appears to be rather ineffective for the group of Austrian participants observed in this study.

A more advanced memory strategy is to combine a word with a visual representation instead of a definition or a translation (Schmitt 1997: 212). Schmitt (1997: 212) stays relatively vague about the benefits of visual representation by stating that it has been demonstrated to be more effective than L1 translation in other languages than English. However, this result has also been validated for the English language by now. Carpenter and Olson (2010: 99) and Seong-Yeon (2007: 27) demonstrated with their studies on 116 students at Iowa State University (Olson 2010: 94) and 70 high-school students from Korea (Seong-Yeon 2007: 27) that pictures had a greater impact on vocabulary learning than translations in English. Demir (2017: 32) investigated whether pictorial representations yielded even better results than sentence examples. However, his hypothesis was falsified. This confirms Carpenter's and Olson's (2010: 92) warning that the effect of the strategy should not be overestimated.

Another memory strategy working with visualisation is semantic mapping. Here a visual framework is made by students between word relationships or meaning relations between words (Schmitt 1997: 213; Nation 2013: 185). This strategy has been reported

to work particularly well with synonyms, antonyms and homonyms of words (Schmitt 1997: 212). However, Gu and Nation (2007: 91) warn that the strategy might have its limitations, since learners might be confused by too many closely related words.

A very frequently studied memory strategy is the keyword technique (Gu & Nation 2007: 90). This strategy requires the learner to “create an unusual association between the word form and its meaning” (Nation 1990: 166). Nation (1990: 166) gives the example of the word *parrot*. The language learners should find a similar sounding word in their L1. In this case he uses the Indonesian word *parit*, which means *ditch*, as an example. The learner is asked to combine the two concepts by, for example, imagining the parrot lying in the ditch. When the learner tries to recall the word *parrot*, the keyword *parit* will work as a clue. While Nation (2013: 465) highlights the immediate effectiveness of the strategy by stating that the keyword technique usually works better than rote learning, use of pictures, imagination, synonyms or guessing from context, he does not forget to mention its limitations. Based on several studies he addresses the controversy that words learned with the keyword technique are forgotten more quickly than words learned with different techniques (Nation 2013: 465). While some studies report that the recall rate after a longer time was high, others claim that results for the long term retention was weaker than for other strategies (Nation 2013: 465).

Schmitt (1997: 213) mentions a similar method to the keyword method, namely the PEG method. To use this strategy a rhyme is memorized. Schmitt (1997: 213) names “one is a bun, two is a shoe, three is a tree” as an example. In a next step, the word that should be learned is combined in a mental image with the word in the rhyme. For instance, if the word to be learned was *chair* the *shoe* mentioned in the rhyme could be imagined standing on the chair. When reciting the rhyme learners will think of the mental images they have created and recall the words learned (Schmitt 1997: 213).

Another vocabulary learning strategy listed is the use of physical action to remember a word. Schmitt (1997: 215) recommends this strategy especially for beginners. However, studies show that multi-sensory (Gorjuan, Hayati & Barazandeh 2012: 349) or tactile strategies (Ogawa et. al. 2014: 5) facilitate language recall and improve long term retention. Gorjuan, Hayati and Barazandeh (2012: 347) conducted a study amongst 60 primary school students, who were tested with the help of several pre- and post-tests on their vocabulary knowledge after learning with visual, tactile, auditory and kinaesthetic

strategies. Results indicated that especially tactile and visual strategies correlated positively with vocabulary learning (Gorjuan, Hayati & Barazandeh 2012: 349). Ogawa et. al. (2014: 73) affirm the results for tactile learning with their qualitative study, using glove-type tactile device, of four male and three female participants.

Nation (2013: 216) has argued that indirect learning can be even more effective than direct vocabulary learning, especially if the latter is poorly planned. What is most important for indirect language learning is the amount and type of input (Nation 2013: 216).

Several possibilities for incidental vocabulary learning are conceivable. Learners could interact consciously with a friend or native speaker in English. This should be done at a regular basis to ensure the right amount of input. Since finding a friend or native speaker might be difficult, language tandems or practice groups at universities might be useful. Furthermore, learners might choose to do translations for professional or vocational purposes. Moreover, students can listen to the English language consciously, watch movies or write correspondences in English. The most studied area of indirect learning is, however, extensive reading. Nation (2013: 218) reports that “research shows that small amounts of incidental vocabulary learning at a range of levels [...] occur from reading. These small amounts can become big amounts if learners read large quantities of comprehensible text”. Comprehensible input means texts which contain only 5 percent or even viewer unknown words as already mentioned earlier (Nation 2013: 218).

2.4.2.2.5 Self-management strategies

After and while applying various consolidation strategies the learners need to use self-management strategies, which allow him/her to monitor his progress, to adapt behaviour and to ensure motivation (Gu & Nation 2007: 92). The self-management strategies used in this taxonomy are a combination of the control strategies mentioned by Gu and Nation (2007: 92) and Hong-Nam's and Leavell's (2006: 404) affective strategies based on Oxford's (1990: 17) definition.

Gu and Nation (2007: 92) summarize Dörnyei's (2001) analysis of self-management strategies by stating that they involve five dimensions of control. Commitment control strategies ensure that learners stay focused towards their goals. Metacognitive control strategies are used by the learners to regulate their concentration. This is linked to

satiation control strategies, which require learners to adapt tasks to make them more exciting and efficient. Environment control strategies are actions that learners take to optimize their learning environment, such as changing space or using different pencils for learning. The fifth strategy named by Gu & Nation (2007: 92) is the emotion control strategy, which is equated with the affection control strategy by Leavell and Hong-Nam (2006: 404) and Oxford (1990: 17). This category involves strategies used to generate positive emotions and to ensure motivation (Leavell & Hong-Nam 2006: 404). Self-management strategies are still a relative new field of investigation in connection with vocabulary learning and will, therefore, need closer examination in the future. In the course of this study the degree of use of self-management strategies of participants will be measured and it will be tried to elicit some concrete strategies applied by learners.

Table 2 presents a summary of all the vocabulary learning strategies considered in this study ordered according to the taxonomy explained.

| |
|---|
| <p>Planning strategies:</p> <ol style="list-style-type: none"> 1. Choosing a word 2. Choosing aspects of word knowledge 3. Choosing strategies 4. Planning repetition |
| <p>Recording strategies:</p> <ol style="list-style-type: none"> 5. Verbal repetition 6. Written repetition 7. Word lists 8. Flash cards 9. Stick sheets on objects 10. Keep a notebook/issue log |
| <p>Discovery</p> <ul style="list-style-type: none"> • <u>Determination strategies (DET):</u> <ol style="list-style-type: none"> 11. Analyse the part of speech in which the word occurs grammatically 12. Analyse affixes and roots of a word 13. Check for L1 cognate or another language 14. Analyse available pictures 15. Guessing 16. Consult a reference source <ul style="list-style-type: none"> Use a monolingual dictionary Use a bilingual dictionary Use an online dictionary Use an existing word list Use existing flash cards Use a corpus |

17. Pass or skip the word
18. Check for defining words or phrases in the text.

- Social strategies (SOC):
 19. Ask the lecturer or an external teacher
 20. Ask classmates or friends
 21. Ask a native speaker

Consolidation:

Direct learning:

- Social strategies (SOC):
 22. Study in a group
 23. Study with a native speaker
 24. Act out role plays
 25. Prepare talks with the vocabulary
- Memory strategies (MEM):
 26. Study words with pictorial representation
 - Label pictures
 - Add a picture to the vocabulary item to visualise it
 27. Imagine a word meaning
 28. Connect word to a personal experience
 29. Use semantic maps
 30. Associate words with coordinates
 31. Connect word to synonyms and antonyms
 32. Use scales for gradable adjectives
 33. PEG Method
 34. Keyword Method
 35. Group words spatially
 36. Use new word in a sentence
 37. Group words within a storyline
 38. Study spelling
 39. Study sound of words
 - Say the words aloud
 - Learn the phonetic transcription
 40. Imagine the word form
 41. Underline initial letter
 42. Configuration
 43. Remember the context in which the word was encountered
 44. Remember word parts
 45. Use cognates for studying
 46. Learn idioms and collocations for a word
 47. Use physical action to remember
 48. Use semantic feature grids
 49. Use pre-existing practice material
 - Sentence completion activities
 50. Use paraphrasing
 51. Use look and recall strategy

| |
|---|
| <p>Indirect learning:</p> <ul style="list-style-type: none"> 52. Interact with a native speaker 53. Interact with a friend in English 54. Doing translations 55. Extensive Reading 56. Glossing 57. Listen to music 58. Watch movies 59. Write private correspondences |
| <p>Self-management strategies:</p> <ul style="list-style-type: none"> 60. Commitment control strategy 61. Metacognitive control strategy 62. Satiation control strategy 63. Affection control strategy 64. Environment control strategy |

Table 2: Vocabulary learning strategies

2.5. Biographical background

The following section will investigate the influence of biographical and personal data on vocabulary learning. While some areas, such as gender or bilingualism, have been well covered in relation to vocabulary learning, other fields of interest, such as a possible relationship between Austrian school types visited by students and their vocabulary proficiency levels at university level have been rather neglected in previous research. Additionally this paper will examine whether experience abroad and English speaking relatives have any influence on vocabulary performance.

2.5.1. Gender

The term gender is used to refer to two related concepts, namely the biological sex of a person and the socially constructed role of this individual (Nyikos 2008: 73). While the differences between biological sexes concerning language learning become more accessible, the influence of gender performativity on language learning is still “neglected as a variable in language learning” (Nyikos 2008: 76). Nevertheless, studies in this field show interesting results, such as parents’ preference to talk more to baby girls than boys. Nyikos (2008: 75) states that “[parents] have longer and more complex conversations with daughters and encourage more responses from them than sons”. Considering this,

gender in both senses can be regarded as influential factor for language learning from the first day on.

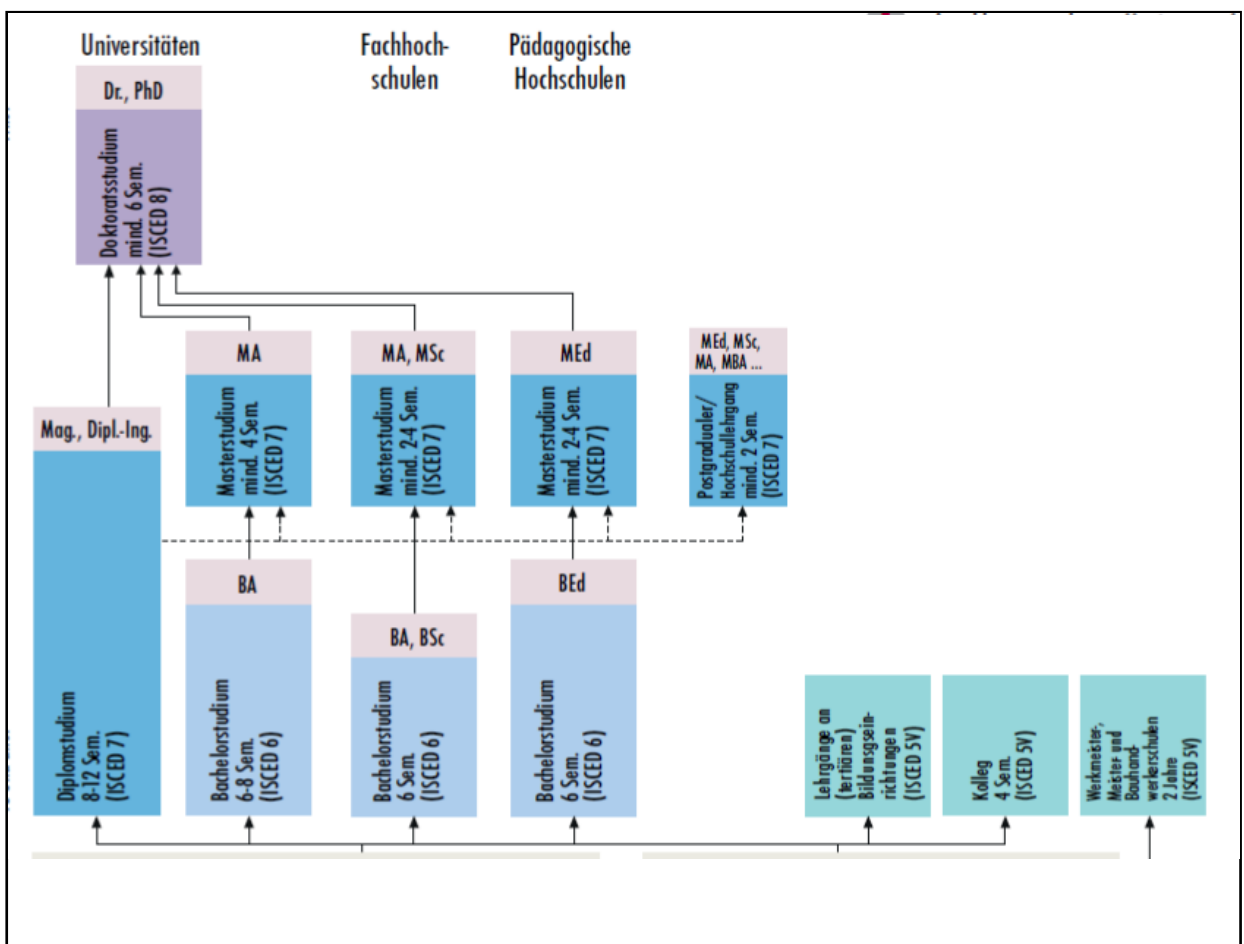
Nyikos (2008: 78) makes some general observations on language learning in relation to sex. She reports that female learners normally attach more value to language learning than male learners. Consequently, they spend more time on language learning as well as on vocabulary learning (Nyikos 2008: 78). Gu (2002) and Catalan (2003) present similar results for the specific field of vocabulary learning strategies. Both studies report that female learners use a greater variety of vocabulary learning strategies than males (Gu 2002: 44; Catalan 2003: 61). Nyikos (2008: 76) adds that females commonly also use more study strategies in general. Additionally to their larger repertoire of vocabulary learning strategies women are more open to trying new approaches to vocabulary learning. Gu (2002: 43), Catalan (2003: 62) and Nyikos (2008: 76-78) further confirm that men and women tend to use different vocabulary learning strategies. Catalan's (2003: 61-62) study of 581 Spanish-speaking participants between the age of 11-56 from all educational levels revealed that there was a statistically significant difference in strategy use, although eight out of ten of the most frequent vocabulary learning strategies were identical for both genders. Female learners appeared to use more formal-rule dependant strategies, more input elicitation strategies and more rehearsal and planning strategies (Catalan 2003: 65). Males, on the contrary, applied more image dependant vocabulary learning strategies, such as forming a mental image of a word (Catalan 2003: 66). Nyikos (2008: 76-78) acknowledges these observations in her study and adds further to them. Women employed more social interactional strategies for vocabulary learning, such as learning in groups. Moreover, they were more successful in employing self-management strategies, since they used self-encouragement strategies, such as rewarding and reminding themselves of their aims (Nyikos 2009: 78). Conversely, men were found to be more goal-oriented and instrumentally motivated than women. Additionally to visual strategies, they employed rote-learning, repetition of all kinds and translation strategies (Nyikos 2008: 78).

Other studies even detect a difference between sexes in vocabulary proficiency reached. Gu (2002: 37-38) investigated adult Chinese EFL learners' vocabulary learning strategies and proficiency levels at the University of Beijing. She utilized a questionnaire on vocabulary learning strategies, the Vocabulary Size Test by Nation and a general proficiency test. Her results indicate that in this group female learners were superior to

male learners, since they outperformed men in both vocabulary size and proficiency (Gu 2002: 40). Nevertheless, other studies (Wyk & Mostert 2016: 1) suggest that gender differences in vocabulary learning are not statistically significant. Consequently, judgments about the relationship between gender and foreign language vocabulary ability should be made with caution.

2.5.2. School type

The school system in Austria leaves open a relatively broad spectrum of choices of school types as can be seen in Figure 13 below, which visualises the Austrian education system from kindergarden to PHD studies at University.



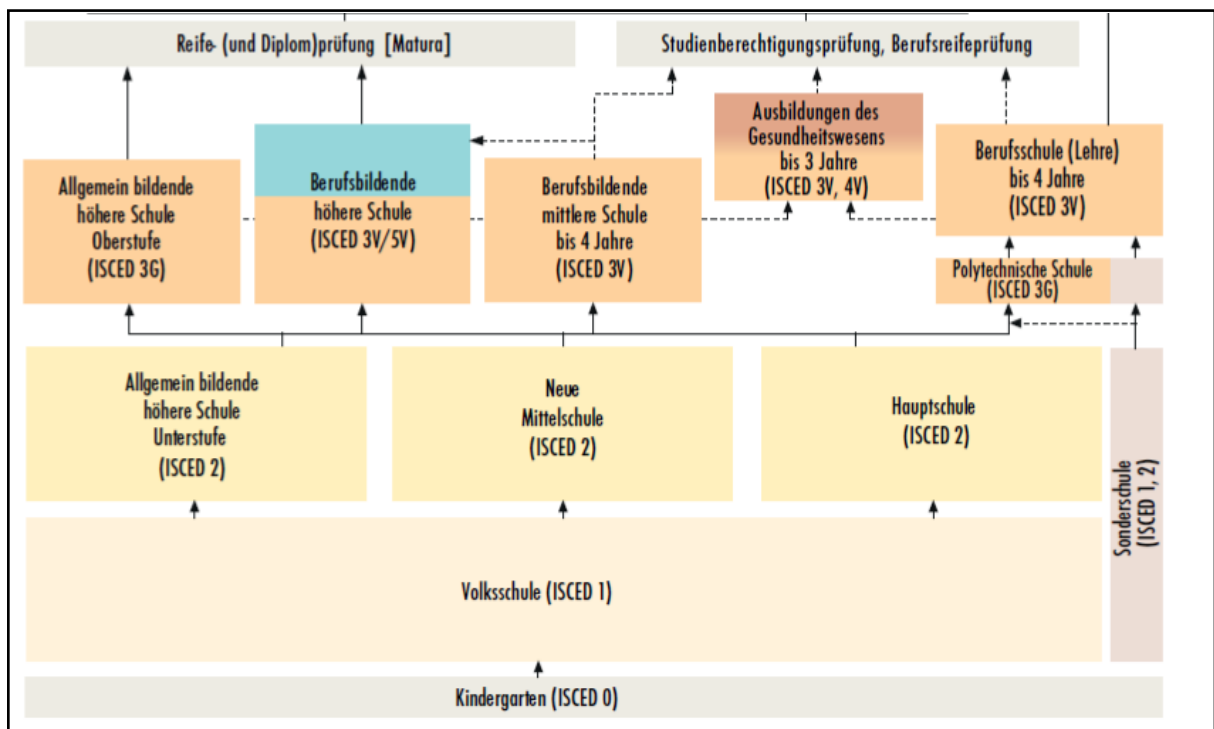


Figure 13: Austrian Education system (Benedik et al. 2017: 17)

The main focus of this study is going to be on the relation between secondary education and tertiary education with a emphasise on university education. As one can see in Figure 13 Austrian students can continue studying at a university after graduating from an **Allgemein bildende höhere Schule** (AHS), which is the equivalent of a grammar school, or an vocational school, named **Berufsbildende höhere Schule** in Austria (BHS). Another possibility is to acquire a Studienberechtigungsprüfung, which is an exam that counts as a general qualification for university entrance. Both ibw (Weiß & Tritscher-Archan 2011: 4) and Statistik Austria (Benedik et al. 2017: 26) report that Austrian students have a strong focus on vocational education with 133,447 students attending a BHS in 2015/16 and only 91,439 students attending an AHS (Benedik et al. 2017: 26) as can be seen in Figure 14.

| Schülerinnen und Schüler nach Schultypen | | | | | | | | |
|---|-----------|------------|-----------|------------|-----------|------------|-----------|------------|
| Schultyp | 1990/91 | | 2000/01 | | 2010/11 | | 2015/16 | |
| | insgesamt | % weiblich | insgesamt | % weiblich | insgesamt | % weiblich | insgesamt | % weiblich |
| Alle Schulen | 1.144.600 | 47,7 | 1.231.188 | 48,7 | 1.166.525 | 48,7 | 1.124.633 | 48,7 |
| Volksschulen | 371.971 | 48,7 | 393.586 | 48,4 | 327.663 | 48,4 | 329.551 | 48,5 |
| Hauptschulen | 238.953 | 48,3 | 263.546 | 47,6 | 192.616 | 47,6 | 28.565 | 48,2 |
| Neue Mittelschulen ¹⁾ | - | - | - | - | 34.324 | 47,2 | 179.480 | 47,2 |
| Sonderschulen | 18.322 | 39,2 | 13.602 | 36,5 | 13.198 | 35,6 | 13.813 | 35,5 |
| Polytechnische Schulen | 19.473 | 30,2 | 19.594 | 34,2 | 18.841 | 37,1 | 16.135 | 35,0 |
| AHS-Unterstufe | 92.878 | 50,1 | 106.925 | 51,8 | 112.330 | 51,7 | 115.631 | 51,9 |
| darunter NMS an AHS ²⁾ | - | - | - | - | - | - | 3.985 | 48,3 |
| AHS-Oberstufe | 65.481 | 54,2 | 77.788 | 57,4 | 88.412 | 56,9 | 91.439 | 57,4 |
| Sonstige allgemein bildende Statutschulen ³⁾ | 1.959 | 49,8 | 5.757 | 48,3 | 9.802 | 49,1 | 9.847 | 48,2 |
| Berufsschulen | 149.806 | 35,2 | 132.613 | 34,0 | 137.881 | 34,5 | 117.358 | 33,7 |
| Berufsbildende mittlere Schulen ⁴⁾ | 67.125 | 59,2 | 64.034 | 60,2 | 78.800 | 59,6 | 73.182 | 59,4 |
| Berufsbildende höhere Schulen | 99.191 | 47,0 | 123.676 | 50,1 | 137.602 | 51,1 | 133.447 | 49,3 |
| Berufsbildende Akademien ⁵⁾ | 2.863 | 78,6 | 4.121 | 83,7 | 1.034 | 83,0 | 111 | 64,0 |
| Lehrerbildende Schulen und Akademien ⁶⁾ | 16.578 | 73,2 | 25.946 | 77,5 | 14.022 | 78,5 | 16.074 | 78,9 |

Figure 14: Distribution of students in school types (Benedik et al. 2017: 25)

Figure 15 visualises how many students of an AHS, BHS or Lehrerbildenden höheren Schule (LHS), which is a school for kindergarten pedagogy, inscribed at an institution of higher education after their school leaving examination. A relatively high percentage of 85.5 of all school leavers of an AHS decided to continue with in any form of tertiary education. Out of these 85.5 percent 85.1 percent chose to attend a public university (Benedik et al. 2017: 62). In comparison, only 54.1 percent of BHS school leavers and 50.9 percent of LHS school leavers decided to pursue tertiary education. Most students out of those who decided for a tertiary education from BHS (69 percent) also decided on a university while 46 percent of LHS school leavers wanting a higher degree tended to choose a Pädagogische Hochschule, which is a college for teacher education (Benedik et al. 2017: 62).

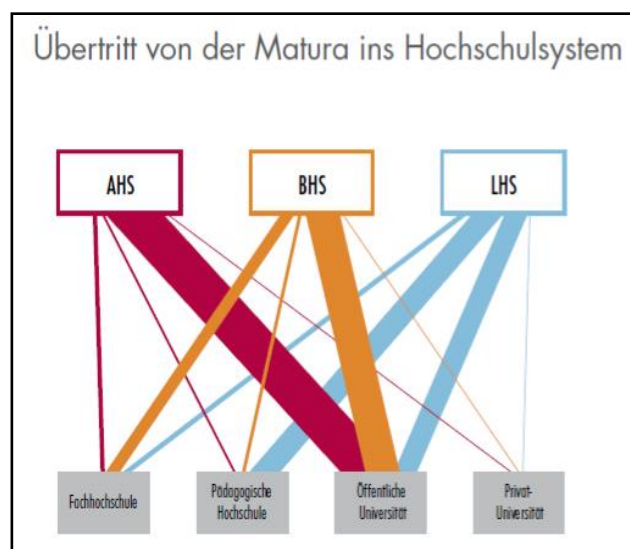


Figure 15: Relationship between school type and tertiary education (Benedik et al. 2017: 63)

2.5.3. Bilingualism

Recent years have brought about an enormous interest in the influence of bilingualism on language acquisition. However, the results are often difficult to compare, since studies base their investigations on different definitions of bilingualism. While some researchers believe that a bilingual needs to be able to use both languages equally well with native-like ability, others take a more liberal approach, defining a bilingual as a person who can utilize two languages to reach certain goals (Keikhaie et al. 2015: 184). Bialystok (2015: 5) brings in a compromise between the two extreme positions by defining bilinguals as “people who are able to speak two (or more) languages, to some level of proficiency”. However, for this study Keikhaie et al.’s view on bilingualism will be adapted, since clear cut boundaries are necessary to deduce results from the data.

Bilingualism has been reported to have an impact on vocabulary acquisition in the L2 and L3 (Keikhaie et al. 2015; Dibaj 2011; Zare & Davoudi Mobarakeh 2013). Dibaj (2011: 193) investigated 52 monolingual Persian speakers and 45 bilinguals speaking Persian and Azeri. All participants were studying English at University level. His results indicate that bilingual speakers clearly outperformed monolinguals in vocabulary acquisition at all levels of learning burden. Similar results were found by Zare and Davoudi Mobarakeh (2013:130) through an investigation of male Arabic-Persian bilinguals high-school students and Persian monolingual high school students studying English, which showed that bilinguals performed better than monolinguals in general and in L3 production vocabulary. However, no difference was found for L3 recognition vocabulary learning (Zare & Davoudi Mobarakeh 2013: 133). Keikhaie et al. (2015: 188) contradict these results in a study of 80 monolingual Persian speakers and 80 Persian-Baluchi bilinguals. While they confirm the general assumption that bilinguals outperform monolinguals in overall vocabulary learning they state that they found no difference for production vocabulary while recognition vocabulary showed a significant effect (Keikhaie et al. 2015: 188). Therefore, they are reporting diverse results than Zare and Davoudi Mobarakeh.

A last study worth mentioning was conducted by Bartolotti and Marian (2016). They investigate how much influence similarities between an L3 and the L1 and L2 have on L3 vocabulary acquisition amongst English-German bilinguals. They conclude that close relationships between languages known and languages learned can have a positive effect on learning success (Bartolotti & Marian 2016: 110). However, inferences between languages might also have a negative impact.

2.5.4. Experience abroad

Increased mobility and flexibility while studying has led to an increased importance of experience abroad, which has sometimes been proposed as “the cure for all language problems” (Kinginger 2011: 58). Although time spent abroad has been praised for improving participants’ language skills by several studies (Kinginger 2011; Munoz 2014) it is crucial to see these universal statements in a more sceptical or critical light. While Kinginger (2011: 59) emphasises the overall positive impact of experience abroad for all types of students, especially for fluency, she highlights the importance of careful planning. In order to benefit, learners need to make sure that they need to engage in local communicative practice (Kinginger 2011: 58). Briggs (2015: 131) relativizes this emphasis on communicative interaction with locals. She reports that the relationship between informal contact and vocabulary gain was not significant in her study. As a possible explanation she argues that students do not always have more native input abroad since they might live alone or socialize in groups of other exchange students with different mother tongues. However, she notes that a positive effect on vocabulary knowledge has been reported by other studies (Briggs 2015: 131). For instance, Ife et al (2000: 55) report that students who had studied abroad all increased their total vocabulary size. Contrary to other studies they observe that there have been no crucial differences in vocabulary gain depending on the proficiency level of the learners (Ife et al. 2000: 55). Llanes and Menoz (2009: 357) add a qualitative dimension to these results by stating that accuracy of vocabulary use has improved amongst their participants through experience abroad. However, here low proficiency students showed more improvement than high proficiency learners. Therefore, one can conclude that students with experience abroad might be likely to have a greater vocabulary size and show more accuracy concerning vocabulary use. However, it is questionable whether this also accounts for academic vocabulary use. This will be explored in this study.

3. RESEARCH

The following section is concerned with the practical part of this paper. Firstly, the research questions will be introduced as well as a description of the participants and the methodology used for the study. These explanations will be followed by the results to all research questions, which will be examined and interpreted more closely in the discussion section. Lastly, a concise conclusion is presented.

3.1. Research questions

The following study aims to discover relations between academic vocabulary proficiency, vocabulary learning strategies and biographical background information. The exact research questions to be answered are as follows:

1. How did academic vocabulary develop from the first academic writing seminar (ILSS1) to the last one (EAP)?
2. How does high-frequency vocabulary and the technical and low-frequency vocabulary subsumed under the “rest” group develop in comparison to academic vocabulary development?
3. Do students show varying proficiency levels concerning academic vocabulary?
4. Is it possible to detect (a) relationship/s between biographical background information and vocabulary learning strategies?
5. Is it possible to detect (a) relationship/s between academic vocabulary levels and vocabulary learning strategies?
6. Is it possible to detect (a) relationship/s between academic vocabulary levels and biographical background information supplied by the participants?
7. How does the accuracy of students with differing academic vocabulary levels vary?

3.2. Participants

The participants in this study are thirty-seven students of the English Department at the University of Vienna. All subjects either had taken or were currently taking the last course in the language competence programme, named English for academic purposes (EAP) and had completed previous courses on academic writing, namely Integrated Language and study skills 1 and 2 (ILSS 1; ILSS 2) amongst others. Therefore, all participants can be considered to be proficient users of the English language.

3.3. Methodology

Participants were asked to complete two parts. Firstly, they had to fill in an online questionnaire on vocabulary learning strategies, which was thoroughly designed according to the guidelines of Dörnyei (2010) amongst others (Chauncey 2013; Meyerhoff, Schleeff & Mackenzie 2015). Secondly, each participant had to provide three uncorrected essays written by themselves, each being from one of the academic writing seminars, ILSS 1, ILSS 2 and EAP, attended during their studies. ILSS 1 is the first course in the language competence programme, while ILSS 2 is in the semester after ILSS 1 and EAP is at the end. Therefore, each essay represents a certain stage in the participants' studies, which makes it possible to visualize a progression or at least a development of vocabulary. Moreover, all three pieces are of roughly the same length, the first two being opinion essays and the last one being an academic book report. This is of crucial importance for reliable results in the corpus analysis conducted later on. All the two essays from ILSS 1 and ILSS 2 were the last assignments in the course and the text from EAP was a book review composed near the middle of the course.

To investigate the percentage of academic vocabulary types and high-frequency types vocabulary used in the corpus the analysis program AntProfiler (2014) was used. This program normally works with the GSL and the AVL. However, it allows users to delete these lists and use others instead. Therefore, in this case the adapted GSL lists including the first 2000 and the 2000-4000 most frequent words and the AVL were used instead as was explained in more detail in section 2.2.6. For this purpose, all three lists had to be rewritten and formatted according to the program's rules. All results were entered into Excel in a next step and later into the statistics program SPSS (2016) to find correlations and to calculate whether the results were statistically significant.

Since all previous results are concerned with quantity, a qualitative part investigating the accuracy of vocabulary use was conducted. An American native speaker was asked to identify all mistakes related to vocabulary usage. In addition, she highlighted all passages which she felt were unacademic. These results will only be presented as descriptive case studies, since the procedure would have to be repeated to be reliable or to allow a quantitative measure of mistakes.

3.4. Results quantitative study

This section presents the results concerning vocabulary development and any correlations with learning strategies and biographical data. All of the research questions will be answered in this section. However, possible reasons for the results will only be discussed in the section 3.6.

3.4.1. Vocabulary development between the different points of measurement

The first research question to be answered concerns the development of academic vocabulary and the change of high-frequency vocabulary, divided into 1-2000 and 2000-4000 most frequent words. Although students beginning at a University should have already acquired a number of high-frequency words at B2 level it remains to be interesting how the percentage of high-frequency vocabulary use changes from the beginning till the end of their studies.

The vocabulary use of students was measured in percentage with the program AntProfiler (2014). Figure 16 shows the mean use of all three vocabulary types at all three points of measurement (ILSS 1, ILSS 2, EAP).

Table 3: Mean of vocabulary types used by students

| Courses | Mean of the vocabulary usage percentages | | |
|---------|--|---------------------------------------|---------------------|
| | High-frequency vocabulary (1-2000) | High-frequency vocabulary (2000-4000) | Academic vocabulary |
| ILSS 1 | 46.69 | 4.31 | 33.12 |
| ILSS 2 | 42.78 | 4.11 | 33.46 |
| EAP | 30.99 | 2.69 | 39.86 |

It can be seen that both high-frequency vocabulary groups show a slight decrease from ILSS 1 to ILSS 2, while the academic vocabulary increased marginally in this time. Between ILSS 2 and EAP the same trends are visible but even stronger.

These differences were investigated with the help of tests, which demonstrate with which probability the difference is only random or statistically significant. To run a significance test a significance level needs to be chosen which functions as a boundary. If the results are higher than the significance level, then the null hypothesis is valid meaning that no change was significant (see Meyerhoff, Schleeff & MacKenzie 2015: 129). For this study,

the significance level was 5%, since this is a common figure used in significance tests (see Bortz & Schuster 2010: 101; Meyerhoff, Schleef & MacKenzie 2015: 130).

To investigate the significance of the vocabulary development a single factor variance analysis with repeated measurement was carried out with the program SPSS (2016). This test can be used if the significance of the results of the same participants at different points in time should be investigated. To use a variance analysis with repeated measurement, it is crucial to examine whether the variances in the individual factor levels and the correlation between the factor levels are homogeneous. This is called the sphericity assumption, which has been inspected with the help of the Mauchly Test of Sphericity (Grinden 1992).

For the academic vocabulary the preconditions for the variance analysis were given as can be seen in Table 4 ($p = .574$). In order to interpret the results of a significance test one needs to look at the p-value, which “represents the probability that our findings are due to chance” (Meyerhoff, Schleef & MacKenzie 2015: 129). Moreover, the count of F is given, since this is common practice in statistical analysis. The results in Table 5 show that the change of the academic vocabulary was highly significant statistically ($F(2, 70) = 27.239$, $p = .000$). Moreover, taking a closer look at the partial η^2 , which is a feature for the effect size of an event depicting whether a circumstance is not only statistically relevant but also practically, it can be seen that it is relatively high. There is no clear boundary value, which specifies, which percentage can be considered as high, middle or low, since the figure is strongly dependent on the research subject. A low η^2 simply means that there are many confounding variables involved. For this study 44 percent in the variation of academic vocabulary use can be attributed to the time of measurement, which can be considered as a medium result.

Table 4: Mauchly’s Sphericity test academic vocabulary

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

| Within Subjects Effect | Mauchly's W | Approx. Chi-Square | df | Sig. | Epsilon ^a | | |
|-------------------------|-------------|--------------------|----|------|----------------------|-------------|-------------|
| | | | | | Greenhouse-Geisser | Huynh-Feldt | Lower-bound |
| Testzeitpunkte_academic | ,968 | 1,110 | 2 | ,574 | ,969 | 1,000 | ,500 |

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept
Within Subjects Design: Testzeitpunkte_academic

Table 5: Significance test for academic vocabulary

| F | Sig. (p) | Partial Eta Squared |
|----------|-----------------|----------------------------|
| 27.239 | .000 | .438 |

Figure 16 visualises the development of academic vocabulary use from ILSS 1 over ILSS 2 to EAP.

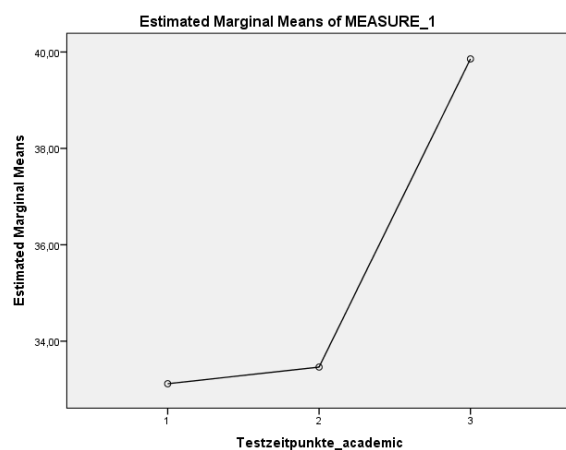


Figure 16: Development of academic vocabulary

Pairwise comparisons in Table 6 show that the differences between ILSS 1 and EAP and ILSS 2 and EAP are highly-significant ($p=.000$) but the development between ILSS 1 and ILSS 2 is not ($p>.05$).

Table 6: Pairwise comparison of academic vocabulary

| Academic vocabulary development Times of measurement compared | Sig. (p) |
|--|-----------------|
| ILSS 1 and ILSS 2 | 1.000 |
| ILSS 1 and EAP | .000 |
| ILSS 2 and ILSS 1 | 1.000 |
| ILSS 2 and EAP | .000 |
| EAP and ILSS 1 | .000 |
| EAP and ILSS 2 | .000 |

As can be seen in Table 7, the Sphericity test for the two groups of high-frequency vocabulary was significant ($p= .043$), meaning that the pre-conditions for the variance analysis are not given.

Table 7: Mauchly's Test of Sphericity high-frequency (1-2000)

Mauchly's Test of Sphericity^a

Measure: MEASURE_1

| Within Subjects Effect | Mauchly's W | Approx. Chi-Square | df | Sig. | Epsilon ^b | | |
|---------------------------------|-------------|--------------------|----|------|----------------------|-------------|-------------|
| | | | | | Greenhouse-Geisser | Huynh-Feldt | Lower-bound |
| Testzeitpunkte_high_freq_1_2000 | ,836 | 6,284 | 2 | ,043 | ,859 | ,898 | ,500 |

As a solution, the Huynh-Feldt corrector was used. Table 8 depicts the results of the significance test for the use of high-frequency vocabulary (1-2000) for each of the three measurements.

Table 8: Significance test for high-frequency vocabulary development (1-2000)

| F | Sig. (p) | Partial Eta Squared |
|----------|-----------------|----------------------------|
| 107.154 | .000 | .749 |

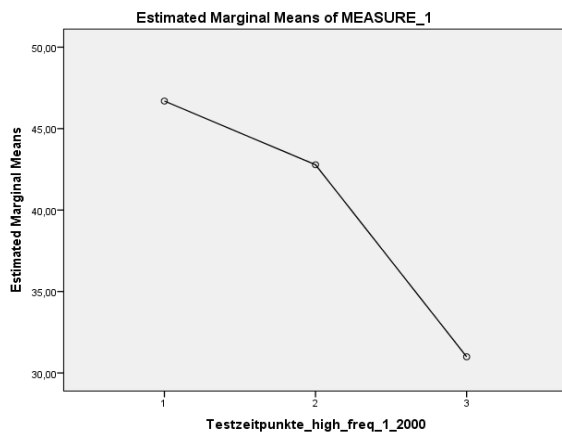


Figure 17: Development of high-frequency vocabulary (1-2000)

It shows that the reduction in use of high-frequency vocabulary, as depicted in Figure 17, is statistically highly relevant ($F= 107.154$, $p = .000$). Taking a closer look at the partial η^2 , it can be seen that the difference of high-frequency vocabulary (1-2000) is practically highly relevant. 75% of the variation in high-frequency vocabulary (1-2000) use is traceable to the time of measurement.

After proving that the change in high-frequency vocabulary (1-2000) was statistically and practically highly relevant, a pairwise comparison in Table 9 shows that the decrease between all three points of measurement are significant ($p < .05$).

Table 9: Pairwise comparison of high-frequency vocabulary (1-2000)

| High-frequency vocabulary (1-2000) Times of measurement compared | Sig. (p) |
|---|----------|
| ILSS 1 and ILSS 2 | .003 |
| ILSS 1 and EAP | .000 |
| ILSS 2 and ILSS 1 | .003 |
| ILSS 2 and EAP | .000 |
| EAP and ILSS 1 | .000 |
| EAP and ILSS 2 | .000 |

Investigating the second group of high-frequency vocabulary (2000-4000) sphericity is again not given and the Huynh-Feldt corrector was applied.

Table 10 shows that the decrease of the high-frequency vocabulary (2000-4000) visualised in Figure 18 is highly significant ($F= 9,133$, $p = .001$). However, the partial Eta² is surprisingly much lower than with the first group of high-frequency vocabulary (1-2000).

Table 10: Significance test for high-frequency vocabulary development (2000-4000)

| F | Sig. (p) | Partial Eta squared |
|-------|----------|---------------------|
| 9.133 | .001 | .202 |

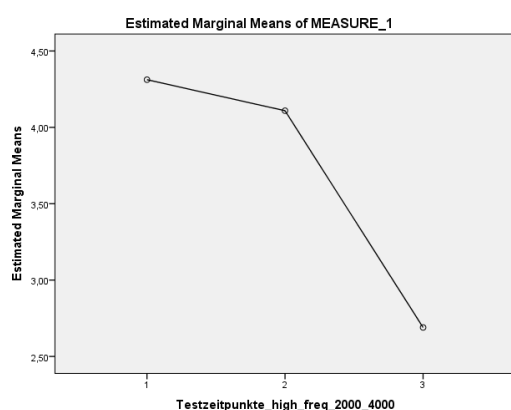


Figure 18: Development of high-frequency vocabulary (2000-4000)

Pairwise comparison shows that only the differences between ILSS 1 and EAP and ILSS 2 and EAP are significant ($p=.000$ bzw. $p=.002$), while the development from ILSS 1 to ILSS 2 was non-significant ($p>.05$) as can be seen in Table 11.

Table 11: Pairwise comparison of high-frequency vocabulary (2000-4000)

| High-frequency vocabulary (2000-4000) Times of measurement compared | Sig. (p) |
|--|-----------------|
| ILSS 1 and ILSS 2 | 1.000 |
| ILSS 1 and EAP | .000 |
| ILSS 2 and ILSS 1 | 1.000 |
| ILSS 2 and EAP | .002 |
| EAP and ILSS 1 | .000 |
| EAP and ILSS 2 | .002 |

The sphericity test for low-frequency and technical words subsumed in one group was not significant, which means that the variance analysis with repeated measurement could be calculated without any correctors as can be seen in Table 12.

Table 12: Sphericity Test for low-frequency and technical vocabulary

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

| Within Subjects Effect | Mauchly's W | Approx. Chi-Square | df | Sig. | Epsilon ^a | | |
|------------------------|-------------|--------------------|----|------|----------------------|-------------|-------------|
| | | | | | Greenhouse-Geisser | Huynh-Feldt | Lower-bound |
| Testzeitpunkte_Rest | ,954 | 1,604 | 2 | ,449 | ,956 | 1,000 | ,500 |

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept
Within Subjects Design: Testzeitpunkte_Rest

The significance test showed that the increase of low-frequency and technical vocabulary as depicted in Figure 19 was highly significant ($F = 67,368$, $p = .000$). Moreover Table 13 shows that, the η^2 is 66 percent, which indicates that the development has a strong practical relevance as well.

Table 13: Significance test for low-frequency and technical vocabulary

| F | Sig. (p) | Partial Eta squared |
|----------|-----------------|----------------------------|
| 67,368 | .000 | .658 |

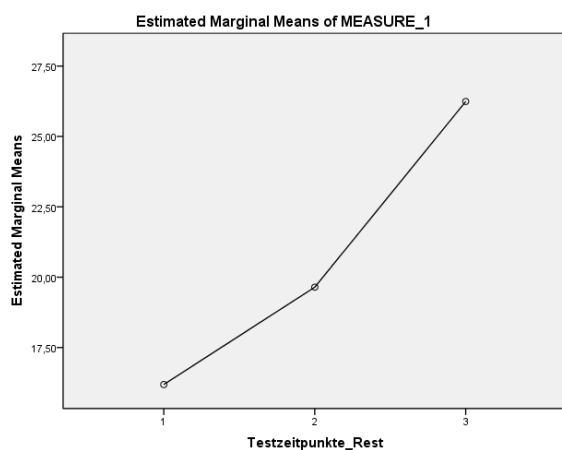


Figure 19 Development of low-frequency and technical vocabulary

Pairwise comparison in Table 14 shows that the development between all courses was highly significant ($p=.000$).

Table 14: Pairwise comparison low-frequency and technical vocabulary

| Low-frequency and technical vocabulary Times of measurement compared | Sig. (p) |
|---|----------|
| ILSS 1 and ILSS 2 | .000 |
| ILSS 1 and EAP | .000 |
| ILSS 2 and ILSS 1 | .000 |
| ILSS 2 and EAP | .000 |
| EAP and ILSS 1 | .000 |
| EAP and ILSS 2 | .000 |

3.4.2. Proficiency levels

After clarifying the development of the different vocabulary types a closer look was taken at the individual results of the participants. All 37 subjects were classified according to their results in one of seven possible groups for each of the three vocabulary categories:

| |
|-------------------|
| 1<2<3 green |
| 1<2>3; 1<3 yellow |
| 1>2<3; 1<3 brown |
| 1<2>3; 1>3 blue |
| 1>2<3; 1>3 violet |
| 1>2>3 red |
| 1<2=3 pink |

Figure 20: Legend of proficiency levels

The green group shows a continuous increase in the vocabulary type investigated, while the yellow and brown groups show an overall increase but had a peak (yellow) or fall (brown) in ILSS 2. The blue and the violet groups had an overall decrease from ILSS 1 to EAP. However, the blue group had the highest result in ILSS 2, while the violet group had its in ILSS 1. The red group depicts continuous decrease. The pink group was only needed in one case, when the result for ILSS 2 and EAP was identical.

Investigating the development of academic vocabulary in Figure 21 it becomes evident that nearly half of the participants (47%) show a continuous increase in academic vocabulary. The second largest group is the brown group (1>2<3;1<3) with 28 percent showing a slight decrease in ILSS 2. This is followed by the violet group (1>2<3; 1>3) with 13 percent. The rest of all participants is evenly spread amongst yellow (1<2>3;1<3), blue (1<2>3; 1>3), red (1>2>3) and pink (1<2=3) accounting for 3 percent. However, one needs to keep in mind that these 3 percent represent exactly one person each.

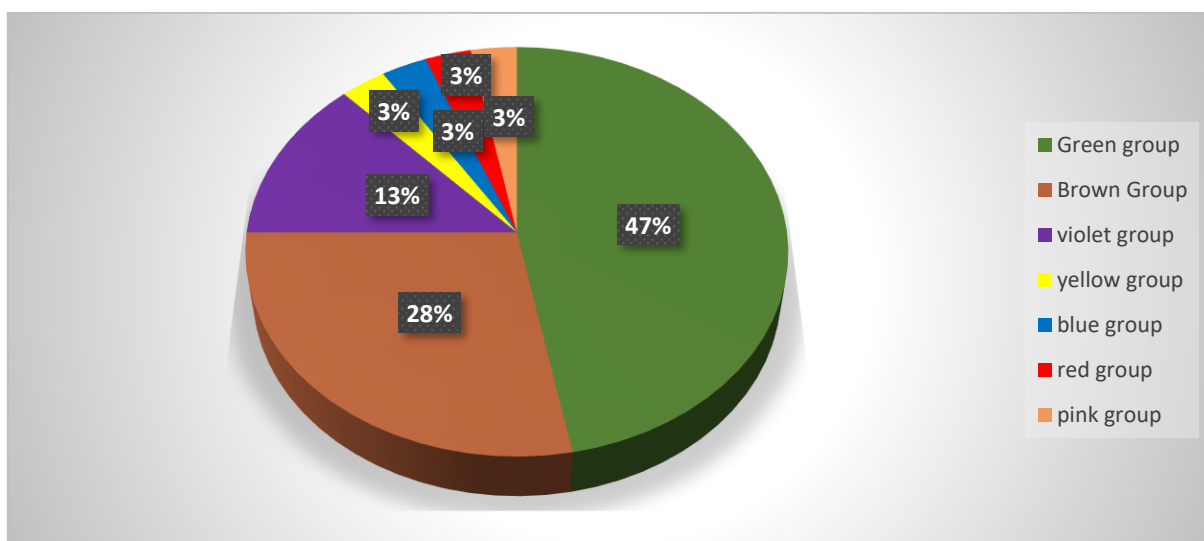


Figure 21: Academic vocabulary proficiency levels

Overall, 81 percent of all participants showed an overall improvement in academic vocabulary and only 19 percent showed a decrease.

Concerning high-frequency vocabulary (1-2000) a reverse picture is visible as expected. In Figure 22 it can be seen that 78 percent show a continuous decrease of high-frequency vocabulary (1-2000), while only 22 percent have a peak in ILSS2 but also show an overall decrease. Nobody used more high-frequency vocabulary (1-2000) in EAP than in ILSS1.

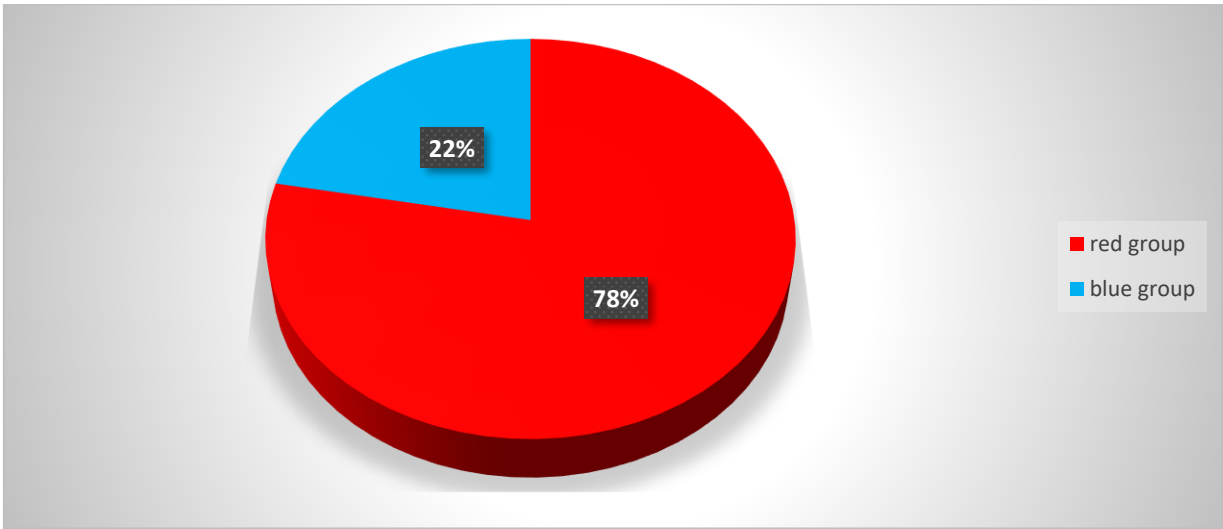


Figure 22: High-frequency vocabulary (1-2000) proficiency levels

Regarding the high-frequency vocabulary (2000-4000) depicted in Figure 23, the majority of participants belongs to the blue (1<2>3; 1>3) or violet (1>2<3; 1>3) group, indicating that most subjects also showed a decrease in the second group of high-frequency vocabulary (2000-4000). 22 percent belong to the yellow group (1<2>3;1<3) followed by 18 percent in the red group (1>2>3) and only 3 percent in the green (1<2<3) or brown group (1>2<3;1<3).

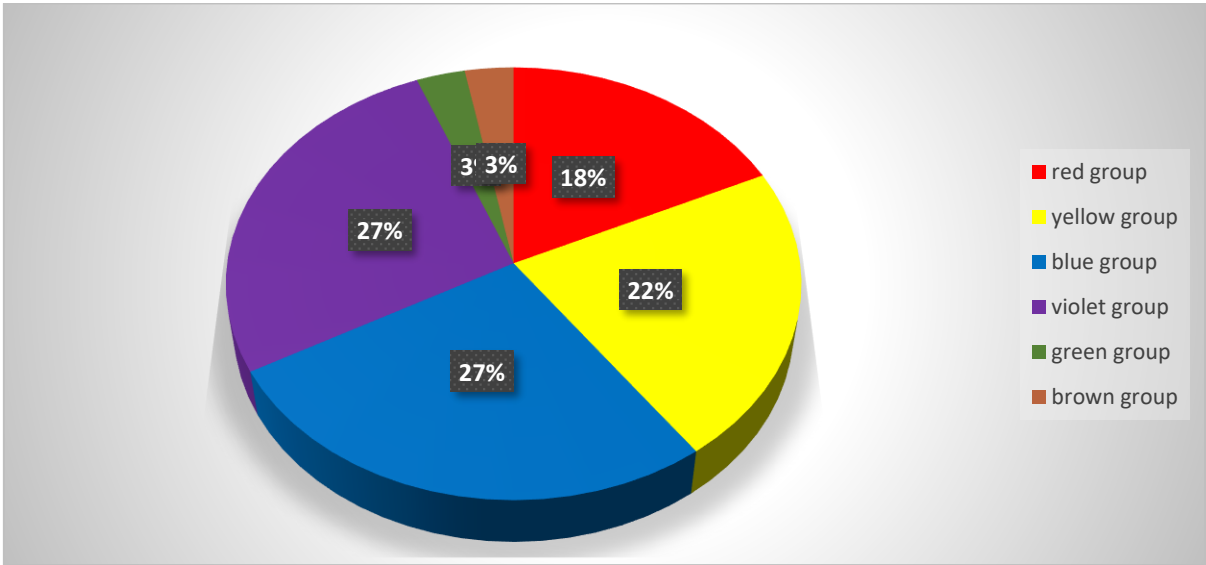


Figure 23: High-frequency vocabulary (2000-4000) proficiency levels

In total, 72 percent used less high-frequency vocabulary (2000-4000) in EAP than ILSS1 and 28 percent used more in EAP than ILSS1.

The results indicate that a higher percentage of academic vocabulary mostly correlates with lower percentages of high-frequency vocabulary in general. However, it can be seen that varying proficiency levels exist, since some participants show atypical results.

This raises the question as to how academic vocabulary changed in the group of participants whose high-frequency vocabulary (2000-4000) use increased instead of decreased. To answer this 10 people were selected. Their high-frequency vocabulary (2000-4000) either increased continuously or between ILSS 1 and EAP overall (green, yellow and brown).

A variance analysis with repeated measurement for these 10 people shows that between ILSS1 and ILSS2 high-frequency vocabulary (2000-4000) increases significantly ($p=.022$) and falls in EAP (however not significantly). The academic vocabulary use of these subjects instead decreases slightly and insignificantly ($p>.05$) between ILSS 1 and ILSS 2 and increases significantly ($p=.001$) by EAP as can be seen in Figure 24.

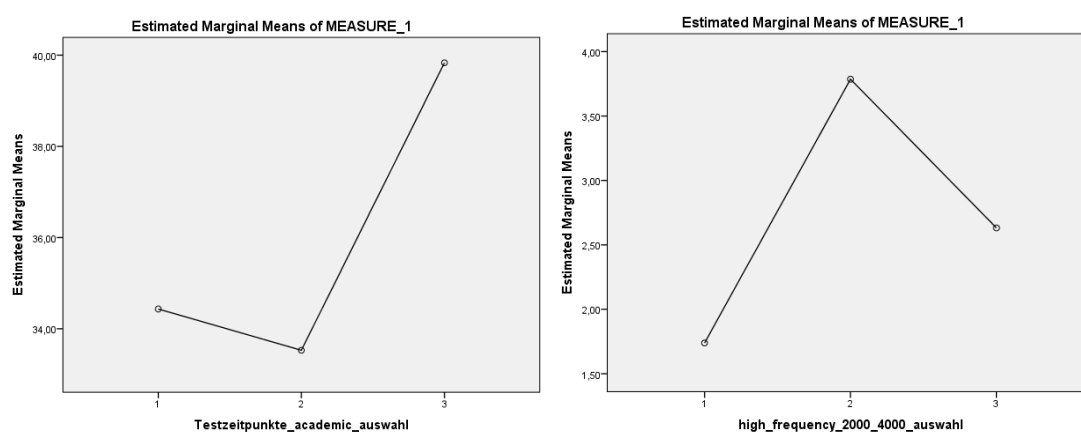


Figure 24: Divergent development of academic and high-frequency vocabulary

Therefore, this analysis confirms the assumption that a negative correlation between academic and high-frequency vocabulary is existent.

3.4.3. Biographical background

In this section research question five will be investigated, namely whether any differences can be detected in the use of vocabulary considering certain context markers. Attention will be paid to gender, school type attended, experience abroad, bilingualism and existence of English speaking relatives.

The gender of the participants of the study was unfortunately not evenly distributed as can be seen in Figure 25.

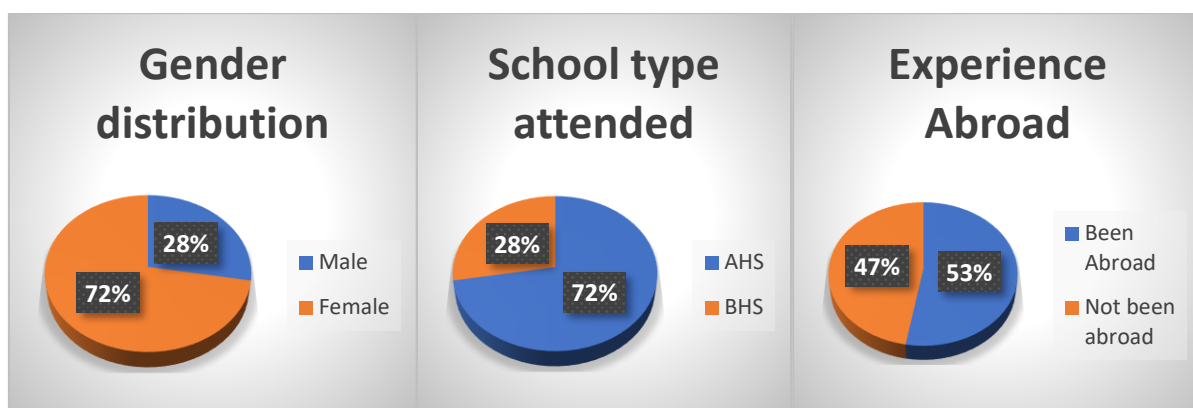


Figure 25: Information on gender, school type and experience abroad

Nearly three quarters of all subjects (72%) were female and only 28% were male. Possible explanations for this uneven result could be that the English appears to be per se a more female field of studies or that female students were more cooperative than male students.

Considering the school type attended most students had attended an AHS (72%) and only 28 percent were graduates of a BHS. Above the exact distribution of the different school types can be seen (see Figure 25).

Regarding participants statements concerning their experience abroad it can be seen in Figure 25 that most subjects had been abroad for some time. While 53 percent had gathered some experience abroad, 47 percent had never been abroad for a longer period of time than some weeks of holidays.

Another interesting point for investigation is the knowledge of other languages that students have. As can be seen in Figure 26 most participants (86%) reported that their mother tongue was German. Only 14 percent indicated that they had another mother tongue.



Figure 26: Information on mother tongue, bilingualism and existence of English relatives

However, a significantly higher number stated that they were bilingual. As Figure 26 visualises 24 percent indicated that they spoke more than two languages at a high level of proficiency. 76 percent of the participants spoken no other language as proficiently than German. Lastly, subjects were asked if they had English speaking relatives. While 28 percent answered with yes, 72 percent negated the question (see Figure 26).

To examine the relationship between these biographical data and vocabulary development, the variance analysis with repeated measurements was supplemented with the listed biographical information. With this method it is possible to detect whether the markers have an influence on the development of vocabulary.

The analysis showed that none of the three different types of vocabulary investigated (academic vocabulary, high-frequency (1-2000) vocabulary, high-frequency (2000-4000) vocabulary) had any significant differences or interaction with any of the biographical information considered. For illustration purposes, the following Figure 27 was added, showing the results for gender and the results for bilingualism.

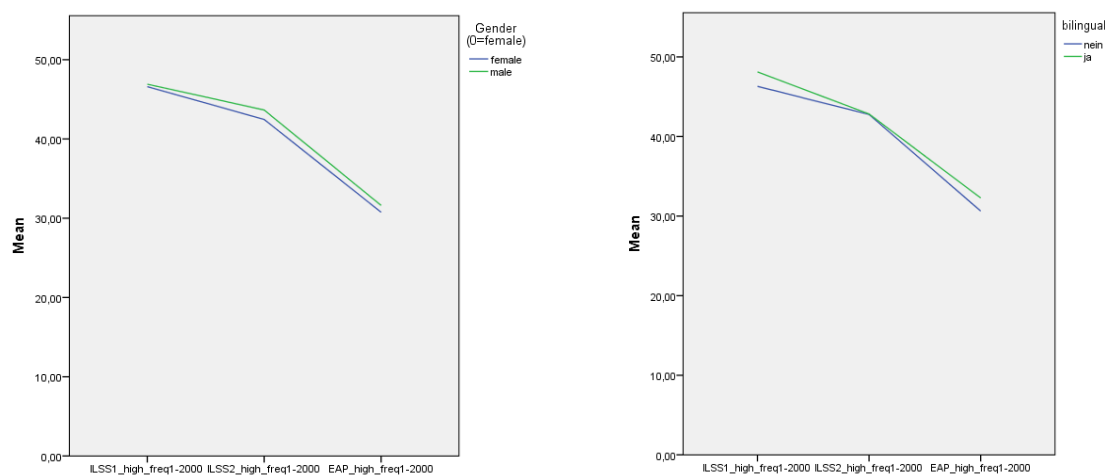


Figure 27: Gender and bilingualism in relation to high-frequency vocabulary (1-2000)

It is visible that there have been no significant differences between the groups. Therefore, biographical background information appears to have no influence on vocabulary development in this case.

3.4.4. Questionnaire

As mentioned in section 3.3.4, a questionnaire was used to investigate which vocabulary learning strategies were used by the participants. This survey was divided into five main categories of vocabulary learning strategies according to the taxonomy in section 2.4.2.3.

- a) Planning vocabulary learning
- b) Recording vocabulary
- c) Discovering the meaning of an unknown word
- d) Learning vocabulary
- e) Self-management strategies

For a closer description of these categories and the strategies subsumed under them see section 2.4.2.3.

The following statistical analysis compares the five categories listed and not the individual strategies, since trend will be more likely to be visible in this way. For this analysis the individual questions of a vocabulary category were transformed into a single scale. The questions for category d) had been answered on a Likert scale with seven possible answers, while all other categories had six possible answers. To make the results comparable all answers being seven (never heard of the strategy) of category d) were

reinterpreted as six (never used the strategy). The content was not altered significantly through this change, since subjects who do not know a strategy cannot use it automatically. Although all questions using a Likert scale were considered for the statistical analysis, answers to open questions will be taken into account in other parts of this results section as descriptive information.

To see how well the internal consistency of the categories is, a scale analysis test was implemented. This test will show how well the feature of a strategy category is captured by a single issue, if all individual questions really measure the same and if a question does not fit in with the others and should be changed or deleted. To be a good example of a category the Cronbach Alpha factor generated by the test should be above 0.60 (see. Bagozzi & Yi; Schnell, Hill & Esser 1999: 47).

As can be seen in Table 15 for the planning strategies the Figure 0.56 was calculated, indicating that the composition of questions might have been more optimal.

Table 15: Cronbach's Alpha

| Planning strategies | Recording strategies | Discovery strategies | Consolidation strategies | Self-management strategies |
|----------------------------|-----------------------------|-----------------------------|---------------------------------|-----------------------------------|
| 0.56 | (0.37); 0.68 | 0.65 | 0.88 | 0.90 |

However, Cronbach's Alpha is dependent on the number of items in a category and can increase with growing number. According to this it can be argued that 0.56 is acceptable, since the category only consists of four questions. The Cronbach's Alpha for discovery, consolidation and self-management strategies are good ranging from 0.65 to 0.90 as can be seen in Table 15.

For recording strategies the first calculation yielded a relatively low figure of 0.37 (see Table 15).

After closer investigation of the inter-item correlation the first question *I do not write down vocabulary in an organised form. I just repeat it orally.* was localised as source for too low Cronbach Alpha (see Table 16 and 17).

Table 16: Inter-item correlation for recording strategies organised vs. disorganised

Inter-Item Correlation Matrix

| | I do not write down vocabulary in an organised form. I just repeat items orally | I do not collect vocabulary in an organised form. I repeat it by writing it down. | I keep a word list. | I make flash cards. | I like to stick sheets with vocabulary written on them on objects in my surroundings | I keep a notebook or vocabulary log. |
|---|---|---|---------------------|---------------------|--|--------------------------------------|
| I do not write down vocabulary in an organised form. I just repeat items orally | 1,000 | ,078 | -,477 | -,270 | -,291 | -,302 |
| I do not collect vocabulary in an organised form. I repeat it by writing it down. | ,078 | 1,000 | ,022 | ,099 | -,019 | ,064 |
| I keep a word list. | -,477 | ,022 | 1,000 | ,617 | ,343 | ,616 |
| I make flash cards. | -,270 | ,099 | ,617 | 1,000 | ,409 | ,302 |
| I like to stick sheets with vocabulary written on them on objects in my surroundings. | -,291 | -,019 | ,343 | ,409 | 1,000 | ,271 |
| I keep a notebook or vocabulary log. | -,302 | ,064 | ,616 | ,302 | ,271 | 1,000 |

Table 17: Item-total statistics for recording strategies

Item-Total Statistics

| | Scale Mean if Item Deleted | Scale Variance if Item Deleted | Corrected Item-Total Correlation | Squared Multiple Correlation | Cronbach's Alpha if Item Deleted |
|---|----------------------------|--------------------------------|----------------------------------|------------------------------|----------------------------------|
| I do not write down vocabulary in an organised form. I just repeat items orally | 21,05 | 26,941 | -,396 | ,256 | ,647 |
| I do not collect vocabulary in an organised form. I repeat it by writing it down. | 21,16 | 19,362 | ,085 | ,031 | ,382 |
| I keep a word list. | 20,68 | 14,781 | ,458 | ,634 | ,126 |
| I make flash cards. | 19,59 | 15,414 | ,492 | ,451 | ,130 |
| I like to stick sheets with vocabulary written on them on objects in my surroundings. | 19,97 | 15,638 | ,266 | ,218 | ,256 |
| I keep a notebook or vocabulary log. | 20,24 | 15,134 | ,383 | ,403 | ,173 |

Selectivity as well as the Inter-item correlation showed that participants having high scores in all other questions had low scores in this question and vice versa. This can be explained by taking a look at the formulation. While in all other questions of the category a high degree of agreement was reflected in a low score in this question agreement was represented by a high score. To level this question the figures of this question were translated (1 interpreted as 6; 5 as 2). A new scale analysis after this change revealed a Cronbach Alpha factor of 0.68, which is again tolerable (see Table 18).

Table 18: Cronbach's Alpha factor for recording strategies second investigation

| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
|------------------|--|------------|
| ,682 | ,685 | 6 |

Therefore, the questionnaire appears to be well designed and should elicit the fitting answers.

3.4.5. Vocabulary learning strategies use and the development of vocabulary use

As can be seen in Table 19 the participants in the study appeared to use most strategies to discover the meaning of an unknown word followed by the strategies to learn vocabulary. It is important to note here that the lowest mean means most use in this case because on the Likert scale 1 meant *used always* and 6 meant *never used*. Therefore, the mean of the discovery strategies (3.2506) and the planning strategies (3.4797) indicates that the participants ticked in average the number 3 standing for *often used*.

Table 19: Mean strategy use

| | Planning strategies | Recording strategies | Discovery strategies | Consolidation strategies | Self-management strategies |
|-------------|----------------------------|-----------------------------|-----------------------------|---------------------------------|-----------------------------------|
| Mean | 3.4797 | 4.0946 | 3.2506 | 4.0923 | 4.6810 |

Consolidation strategies were ranked on the third place, while recording strategies were second to last. The least used strategies were self-management strategies. However, participants named some interesting self-management strategies in the course of learning vocabulary. To remind themselves of their learning goals students “[wrote] them on a sheet of paper”, “[made] a checklist” and “re-read texts with previously unknown words after having looked up the meaning”. Additionally, participants reminded themselves that they would speak more fluently if they knew more vocabulary items, compared themselves to native speakers and reflected on the importance of vocabulary learning for productive skills to remind themselves of the value of learning vocabulary. To regulate their concentration students reported to change place, make short breaks, use different pens and reward themselves. Adaptions named made to avoid boredom were trying to find "fancy" vocabulary items that can be used in their own creative endeavours, drawing, saying the word in a funny way, flash cards, lists, speed rounds of repetition, imagining what the vocabulary describes and doing research on collocations. Students also reported

using relaxation techniques such as making short naps, doing yoga or sports, having a beer, massaging their ears, stretching and meditation. Almost all students giving feedback on affective strategies said that they worked with a rewarding system.

To analyse the relationship between vocabulary learning strategies and the vocabulary use of participants, different approaches were taken. Firstly, the variance analysis with repeated measurement could not be calculated with the figures of the Likert scales (1-6), since this results in too many characteristics. To reduce these participants were first divided into three groups according to the intensity participants used a category of strategies, such as planning strategies or discovery strategies, with (0-2= always/often; 2-4= sometimes; 4-6= rarely/never). However, this proved to be ineffective since the mean calculated for strategies started above 2. Therefore, the participants were divided into two halves next and the analysis was computed. Unfortunately, no relationship between vocabulary strategies used by one of the two groups and their vocabulary development was found. Both groups had a similar academic and high-frequency vocabulary usage. For illustration purposes Figure 28 below shows the results of both groups for the category planning vocabulary learning and high-frequency vocabulary (1-2000).

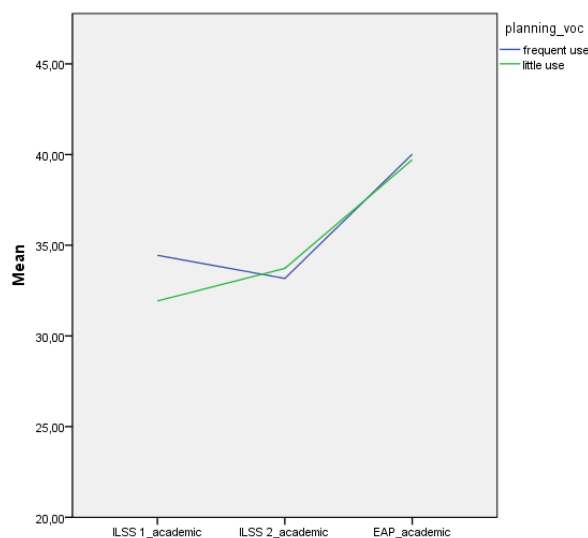


Figure 28: Academic vocabulary and planning strategies

As one can see the second half of participants used these strategies slightly more. However, the difference is not big and not statistically significant.

Secondly, participants were divided into thirds and the top and bottom group were compared, meaning the upper and lower thirds of the subjects, since they could indicate

the potentially largest differences. For this grouping no significant results were generated as can be seen in Figure 29 below illustrating the same example as above, namely the relation between the planning strategies used by the groups and their high-frequency vocabulary (1-2000) development.

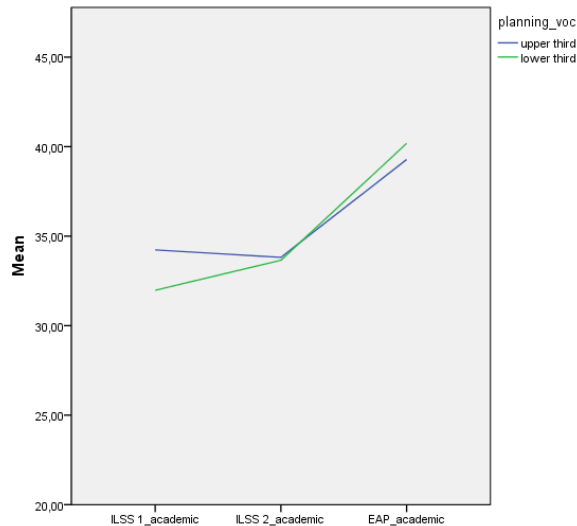


Figure 29: Academic vocabulary and planning strategies

Thirdly, it was discovered that the biggest drop of high-frequency vocabulary (1-2000) can be detected between ILSS 2 and EAP. For each participant the exact margin between the percentages in ILSS 2 and EAP was determined. In the next step a Spearman correlation was calculated to see if the change between these two points of measurement is related to the learning strategies. However, this test also did not generate any statistical relevant positive results. No relation between the difference in vocabulary use and a vocabulary learning strategy could be found.

Lastly, it was analysed if the ten participants (see section 3.4.2) who increased their high-frequency vocabulary (1-2000) instead of reducing it used different learning strategies than other subjects. Again, no significant differences were found.

These findings have resulted in a clear picture, namely that the development of vocabulary types used for academic writing were not related to vocabulary learning strategies in this study. Therefore, research question five was refuted. However, there might be possible explanations for these results, which will be addressed in section 3.5. Moreover, specific trends for the participants in this study can be noted although they are not statistically relevant.

3.4.6. Vocabulary learning strategies and biographical data

Three correlations can be detected when comparing biographical data with the use of vocabulary learning strategies.

Gender proved to cause differences. The mean of men and women clearly show that all strategies are used more often by women as can be seen in Table 20.

Table 20: Gender distribution of strategies

| Gender | Planning strategies | Recording strategies | Discovery strategies | Consolidation strategies | Self-management strategies |
|---------------|----------------------------|-----------------------------|-----------------------------|---------------------------------|-----------------------------------|
| Female | 3.2870 | 3.8642 | 3.1751 | 3.9928 | 4.4147 |
| Male | 4.0000 | 4.7167 | 3.4545 | 4.3611 | 5.4000 |

Both sexes use the strategies to discover the meaning of an unknown word most and self-management strategies least.

To analyse if and how the usage of vocabulary learning strategies depends on gender a t-test was used. It shows that the strategies for planning, recording and self-management were used significantly more often by women than men. For the other two categories no significant results were found as can be seen in Table 21.

Table 21: T-test Gender and vocabulary strategies

| | Planning strategies | Recording strategies | Discovery strategies | Consolidation strategies | Self-management strategies |
|-----------------|----------------------------|-----------------------------|-----------------------------|---------------------------------|-----------------------------------|
| Sig. (p) | .012 | .020 | 1.91 | .092 | .011 |

Moreover, the Cohens d factor, which measures the effect size of a result shows that the differences between male and female use of planning, recording and self-management strategies is highly practically relevant as can be seen in Table 22. Discovery strategies and Consolidation strategies yielded no significant results.

Table 22: Cohens d factor

| | Planning strategies | Recording strategies | Self-management strategies | Consolidation strategies |
|-----------------|----------------------------|-----------------------------|-----------------------------------|---------------------------------|
| Cohens d | 1.026 | 1.019 | 1.145 | 0.89 |

The Cohens d factor is said to be big if it is higher than 0.8. In this case the results even reach 1.026, 1.019 and 1.145. Respectively, self-management generating the best result and planning strategies the lowest of the three.

Another slightly significant difference in the use of strategies can be found between the group with English speaking relatives and without English relatives. Here the strategies subsumed in the category vocabulary learning strategies were used more often by subjects with English relatives as can be seen in Table 23. Moreover, the t-test for this correlation was positive ($p=.025$).

Table 23: Mean of vocabulary learning strategies and English speaking relatives

| English speaking relatives | Mean |
|-----------------------------------|-------------|
| No | 4.2366 |
| Yes | 3.7028 |

In addition, self-management strategy use differed according to the school type attended. Participants stating that they had finished an AHS-Oberstufe (grammar school) used slightly more self-management strategies (4.5767) than BHS (vocational school) school leavers (5.1806) as depicted in Table 24.

Table 24: Mean of Self-management strategies and school type

| School type visited | Mean |
|----------------------------|-------------|
| BHS | 5.1806 |
| AHS | 4.5767 |

In addition, the t-test showed that this correlation was highly significant ($p = .036$).

Table 25: Group statistics school type

| Group Statistics | | | | | |
|----------------------------------|---|----|--------|----------------|-----------------|
| | Abschluss Sekundarstufe 2 (1 = berufsbildende, ...) | N | Mean | Std. Deviation | Std. Error Mean |
| Skala self-management strategies | berufsbildende höhere Schule (HLW, HAK, HTL) | 9 | 5,1806 | ,48858 | ,16286 |
| | AHS Oberstufe | 27 | 4,5767 | 1,15405 | ,22210 |

However, the results of the comparison of the factors English speaking relatives and school type were relatively small and no other differences for any of the biographical factors could be detected.

3.5. Results qualitative study

Since the major part of this study is concerned with quantitative results, this investigation was carried out to pay attention to the qualitative side of the study. A native speaker was asked to mark all mistakes concerning vocabulary in the texts and to highlight all formulations that she considered to be unacademic.



Figure 30: Legend of proficiency levels

Taking a look at the mistakes it is striking that all texts have a relatively low number of mistakes in general. However, participants with no overall improvement of proficiency level in academic vocabulary (red, blue and violet groups) did not differ strongly from subjects with overall improvement (green, brown and yellow groups). For both groups the number of mistakes decreased from ILSS 1 to EAP in most participants. In the last seminar most subjects had no mistakes at all. There were only two exceptions. The participant showing a continuous decrease in academic vocabulary (red group) had more mistakes in EAP (“small languages”, “standing strong” “great overlook”, “only one speaker”) than in ILSS 1 (“long time period”). Furthermore, the subjects in the brown group who had a lower percentage of academic vocabulary in ILSS 2 than in ILSS 1 also had more mistakes in their ILSS 2 essays than in the other two (“closely bonded”, “freshest

number” “rethought statement”, “once body”, “fierce desire”, “interventions into the body”). Therefore, the development of the vocabulary mistakes appears to reflect the overall academic vocabulary development.

Another interesting point to note is that most mistakes made were due to the incorrect use of a word (“Parents should be enlightened about the importance”, “knowledge has to be transmitted”, “discharge environment”) or due to the incorrect use of a collocation (“to know by hart”, “great overlook”, “stand firm”). Since students were able to reduce these mistakes over the time, it could be argued that in most cases the mistakes can be seen as a sign of progress and change in the interlanguage of participants (Lightbrown & Spada 2013: 43). Only one participant appears to have fossilized (Lightbrown & Spada 2013: 43).

Regarding the parts classified as being unacademic a similar trend to the mistakes is visible. Most participants started with a larger number of informal expressions (“constantly staring at a mobile phone”, “put the phone away”, “surely a good thing”, “frankly speaking”, “make everything different”, “a big number of”, “a lot of influences”) in ILSS 1 and ended with none in EAP. This highlights the strong progress made. Nevertheless, also this trend has exceptions. The number of informal statements did not always correspond to the proficiency level in academic vocabulary. For instance, one member of the green group showing continuous progress in academic vocabulary had a similarly large number of unacademic terms in the last essay as in the first. Additionally, the brown groups showed a higher level of informality in ILSS 2 than in ILSS 1 and EAP as they already did before with the mistakes.

To sum up, the development of mistakes and formality of expression appears to conform to the overall development of academic vocabulary. Therefore, one can state that students did not only improve concerning quantitative measurements but were also able to improve their accuracy and quality throughout their studies.

3.6. Discussion

In this part of the paper the results presented above will be discussed, compared and suggestions for possible explanations will be given.

Research question one which is concerned with the development of academic vocabulary from ILSS 1 to EAP has been answered satisfactorily. Participants showed a highly significant statistical and practical increase in academic vocabulary from ILSS 1 to EAP. Therefore, the overall academic vocabulary development of students in the academic setting of the English Department appears to be successful since students were able to improve their performance. This means that they have become more competent and flexible members of the academic discourse community, which can be seen as one of the ultimate goals of university education in general.

However, it has been noted before that the development of academic vocabulary was only significant between ILSS 1 and EAP and ILSS 2 and EAP but not between ILSS 1 and ILSS 2 as can be seen in section 3.4.1.. This result is easily explained by the average time passing between the courses. While ILSS 1 and ILSS 2 can be taken in two successive semesters, the minimum time between ILSS 2 and EAP is much bigger. Three other courses must be completed positively before EAP can be attempted. This leaves a minimum time of one and a half years between ILSS 2 and EAP. According to this, students had more time to improve their academic vocabulary from ILSS 2 to EAP than from ILSS 1 to ILSS 2, which could explain why only ILSS 2 to EAP and ILSS 1 to EAP is statistically significant. In addition, ILSS 1 and ILSS 2 are two of the first courses in the studies. Therefore, students might have needed some time to adapt to the academic setting and its requirements, while at a later stage in EAP students might be able to work more effectively already and those who could not adapt had dropped out.

Another interesting question arises concerning the results to research question two, which asks for the development of high-frequency vocabulary in comparison to academic vocabulary. It is visible that both groups of high-frequency vocabulary (1-2000 and 2000-4000) are indirectly proportional to academic vocabulary. If the academic vocabulary increases high-frequency groups decrease. Even in the subjects who showed an atypical use of academic vocabulary with a decrease in parts of their studies, the same relationship between academic and high-frequency vocabulary was evident by the end of EAP. This

result is not as unsurprising as it sounds, however. It might appear logical that if a text consists of 100 percent of words the percentage of the second vocabulary type must fall if the other group increases. Nevertheless, one must not forget that an academic text does not only consist of high-frequency and academic vocabulary, but also contains low-frequency and technical vocabulary considered as “the rest” in this study. Hence, the relationship between high-frequency vocabulary and academic vocabulary is not as straight-forward anymore, since technical and low-frequency vocabulary increased statistically significantly similar to academic vocabulary and only high-frequency vocabulary shows a decrease.

Another point worth mentioning at this stage of the discussion is the existence of varying proficiency levels. This research question has been answered positively, since participants showed different vocabulary development, although the mean continuously increased for academic vocabulary and decreased for high-frequency vocabulary. Especially the difference between the first group of high-frequency vocabulary (1-2000) and the second group (2000-4000) is outstanding. While the first group of high-frequency vocabulary (1-2000) is only divided into two groups who developed differently (see section 3.4.2.), the second group is split in six subgroups. Based on this result one might consider that high-frequency vocabulary is not behaving homogeneously and that the two groups behave differently concerning development. For future research it might be interesting therefore to split high-frequency vocabulary in even more than two groups to see if a boundary can be detected where high-frequency vocabulary appears to behave differently in relation to other vocabulary types in academic writing.

The most surprising results in this study were that vocabulary development did not correlate positively with biographical data (research question four) or certain vocabulary learning strategies (research question five).

Possible reasons for the lack of correlation with biographical background information could be that influences triggered by different biographical starting conditions were compensated by the input at university. This might also explain different percentages of academic vocabulary in ILSS 1. Some students already started with a relatively high percentage of academic vocabulary, while others were not as successful. However, over the time of the complete studies these differences due to preconditions might decrease. Other influential factors could have been IQ, aptitude or degree of work enthusiasm.

The missing influence of some biographical factors can also be explained by different influences. Concerning English speaking relatives it must be noted that most students who had indicated that they had English speaking relatives reported that they had only sporadic contact: "Yes but I seldomly speak to them.", "Yes, a few times per year.", "Yes but not so often in contact". Therefore, the different degrees of intensity of contact might have been too diverse to be summarized under yes or no. Moreover, participants will most likely not use academic vocabulary in conversations with their relatives either.

Regarding the factor of experience abroad it is interesting that many participants had been abroad but not in an English speaking country ("Belgium 2017, 5 months", "Valencia 2016/17 5 months") or at least not at an University ("London, 2013-2017, altogether around 4 months (NOT CONTINUOUSLY THOUGH - I went there every other month for a few days/weeks to visit my boyfriend and his family)"). Therefore, it can be argued that the influence of experience abroad might not have been as meaningful as expected, since visiting friends or working as an Au-pair might be beneficial for fluency or other abilities but is likely to have no crucial influence on the development of academic vocabulary. To prove whether studying abroad has an influence, the study would need to be repeated with more carefully selected participants for this purpose. However, at this stage this study contradicts the results of Ife et al. (2000) and Llanes and Menoz (2009), since participants with any experience abroad did not show a larger vocabulary repertoire or a higher level of accuracy.

The non-existing relationship of vocabulary development and gender and bilingualism is, however, surprising. Referring back to the controversy detected between Keikahaie et al.'s (2015) and Zare and Mebarakeh's (2013) studies on the influence of bilingualism on vocabulary learning in section 2.5.2 it could be noted that the results of this study might support Keikhaie et al.'s (2015) results, since they reported no significant difference between monolinguals and bilinguals for vocabulary production. Moreover, Bartolotti and Marian (2016: 112) name differing degrees of similarities between languages as a possible reason for non-existing or divergent results on vocabulary influence. Therefore, it could be investigated further if the languages spoken by the subjects (French, Serbian, Hungarian...) have similarities to English, such as cognate relations, or if the relationship was not strong. Moreover, it would be crucial to survey how bilingualism was defined by the participants to see if there were differences.

Research question five asking for the relationship between vocabulary learning strategies and vocabulary development had to be refuted as well. Therefore, no relationship could be found. One possible explanation could be that all students, even those with a low proficiency level in academic vocabulary, used vocabulary strategies. However, the questionnaire makes no statement about the effectiveness with which students apply these strategies. Since during their studies at the English Department of the University of Vienna keeping a vocabulary issue log is a requirement, students have to try out new methods. Nevertheless, it is not controlled if students use the strategies effectively or even attempt to learn the vocabulary collected. Therefore, a possible explanation for the non-existing relationship between vocabulary learning strategies and vocabulary development, as well as varying proficiency levels could be that more proficient students used the strategies successfully, while less proficient students might use the same strategies but not as effectively. This could be due to lacking commitment or that these students are not able to transfer their new passive, receptive knowledge to their active productive knowledge as well as others.

Additionally, it is surprising that the mean of most vocabulary learning strategies is higher than two, which means that students tried to avoid extreme statements. A possible reason could be that participants do not want to admit that they may use only some strategies extensively instead of a broader variety against their better knowledge. Therefore, the information gathered by the questionnaire might not be completely reliable. For future research projects it might be more sufficient to group participants and instruct them to learn only with one specific vocabulary strategy for a specific time before examining their vocabulary use. For instance, one group could consist of students learning mainly with the help of word lists. This approach would ensure that all participants in a group had really used the approach intensively and, if the quality of learning was controlled, also effectively. A repeated measurement of vocabulary quantity would maybe result in divergent results to the recent ones.

The relationship between vocabulary learning strategies and biographical information revealed mostly corresponding results with previous studies (Gu 2002; Catalan 2003, Nyikos 2008). Similar to Gu (2002: 44) and Catalan (2003: 61) this research revealed that women use more vocabulary learning strategies than men. It was also evident that female subjects used different strategies than their male counterparts. Women invested more in planning and recording vocabulary, which validates Catalan's (2003: 65) observations,

who reported that her female participants used more planning strategies. However, the dominance of women in the use of recording strategies could be interpreted as a contradiction of the given literature (Nyikos 2008: 78). So far males were found to be using more visually oriented strategies than women.

Additionally, the study revealed that women used significantly more self-management strategies than men, which validates already existing results (Nyikos 2008: 78). However, women did not learn in a more social or interactive manner as suggested by Nyikos (2008: 76-78). What is even more surprising is that the findings concerning gender dependant strategy use is aligning to previous findings but that in this study women did not outperform men in vocabulary size or accuracy. Gu (2002: 40) stated that female strategy use yielded better results in these areas. This was not confirmed in this study. Women showed most typical signs in their strategy use but these did not lead to better results than male strategy use did.

The different strategies use noted for participants with or without English speaking relatives and from different school types were relatively small but statistically significant. Therefore, further investigation might be rewarding.

4. Conclusion

To sum up, this study has shown that students at the English Department of the University of Vienna increased their use of academic vocabulary from the beginning till the end of their studies statistically and practically significantly. However, participants showed varying proficiency levels concerning academic vocabulary. High-frequency vocabulary correlated negatively with academic vocabulary and decreased from measurement to measurement. However, it was interesting to see that the less frequent group of high-frequency vocabulary (2000-4000) investigated showed a more diverse picture of varying proficiency levels than the first 2000 high-frequency vocabulary items. Therefore, it might be interesting to see if a boundary for this mixed behaviour can be found. In contrast to high-frequency vocabulary, low-frequency and technical vocabulary increased from each course to the next. Qualitative results corresponded to the academic vocabulary development, since accuracy and formality increased throughout the studies.

No relationship could be found for vocabulary development and biographical data. Nevertheless, future investigations might reveal divergent results if the choice of participants was more accurate for the biographical aspect to be measured. A similar result was found for vocabulary development and vocabulary learning strategies. No relationship could be detected. Nonetheless, it would be important for future studies to focus on the effectiveness of the strategies applied and not solely on quantitative measures. Maybe this would yield different results.

Lastly, a connection between vocabulary learning strategies and the biographical factors gender, English relatives and school type attended was found. While the results on gender validate previous findings of other researchers, the other two findings open up new perspectives and might therefore be worth the effort of further investigation.

Word count: 27287

5. References

- Alabsi, Thuraya. 2016. "The effectiveness of role play strategy in teaching vocabulary". *Theory and Practice in Language Studies*, 6 (2), 227-234.
- Ansarin, Akbar Ali; Zohrabi, Mohammad; Zeynali, Simin. 2012. "Language Learning strategies and vocabulary size in Iranian EFL learners." *Theory and Practice in Language Studies* 2(9), 1841-1848.
- Anthony, Laurence. 2014. *AntProfiler*. (Version Windows 1.4.0). [Computer Program] <http://www.laurenceanthony.net/software.html> (14 July 2017).
- Anthony, Laurence. 2016. *AntConc*. (Version Windwos 3.4.4). [Computer Program] <http://www.laurenceanthony.net/software.html> (14 July 2017).
- Aronin, Sara; Haynes-Smith, Heather. 2013. "Increasing science vocabulary using Powerpoint flashcards". *Science Scope* 37(3), 33-36.
- Atkins, Beryl; Varantola, Krista. 1997. "Monitoring dictionary use". *International Journal of Lexicography* 10(1), 1-44.
- Austin, John. 2010. *How to do things with words*. Stuttgart: Reclam.
- Bagozzi, Richard; Yi, Youjae. 1988. "On the Evaluation of Structural Equation Models". *Journal of the Academy of Marketing Science* 16(1), 74-94.
- Barcroft, Joe. 2009. "Strategies and performance in intentional L2 vocabulary learning." *Language Awareness* 18(1), 74-89.
- Briggs, Jessica. 2015. "Out-of-class language contact and vocabulary gain in a study abroad context". *System* 53, 129-140.
- Bartolotti, James; Marian, Viorica. 2016. "Bilinguals' existing languages benefit vocabulary learning in a third language". *Language Learning* 67(1), 110-140.
- Beauvoir, Simone de. 1949. *The Second Sex*. London: Vintage Books, 295-351.
- Benedik, Oliver Mag. Markus Bönisch; Harald Gumpoldsberger; Mag. Sarah Klem; Sabine Martinschitz; DI Dr. Gerhard Nachtmann; Mag. Wolfgang Pauli; Mag. Jakob Peterbauer; Mag. Regina Radinger; Mag. Manuel Reif; Mag. Natascha Riha; Mag. Brigitte Salfinger-Pilz; MMag. Guido Sommer-Binder; Mag. Cornelia Speckle; Klaus Trenkwalder MA; Mag. Barbara Wanek-Zajic; Mag. Alexander Wisbauer. 2017. "Bildung in Zahlen 2015/16 [Education in numbers 2015/16]". *Statistik Austria*. <http://www.uniko.ac.at> (5. Aug. 2017).
- Biemiller, Andrew; Boote, Catherine. 2006. "An effective method for building meaning vocabulary in the primary grades". *Journal of Educational Psychology* 98(1), 44-62.
- Bogaards, Paul; Laufer, Batia. 2004. *Vocabulary in a Second Language*. Amsterdam: John Benjamins.
- Boonmoh, Atipat. 2012. "E-dictionary use under the spotlight. Students' use of pocket electronic dictionaries for writing". *Lexikos* 22, 43-68.
- Bortz, Jürgen; Schuster, Christoph. 2010. *Statistik für Human- und Sozialwissenschaftler*. Heidelberg: Springer.
- Bradley, Lynette.; Huxford, Laura. 1994. "Organising sound and letter patterns for spelling". In Brown, Gordon; Ellis, Nick. *Handbook of spelling*. Chichester: John Wiley and Sons, 425-439.
- Brosch, Cyril. 2015. "On the conceptual history of the term lingua franca" *Journal of Applied Language Studies* 9(1), 71,85.
- Browne, Charles; Culligan, Brent; Phillips, Joseph. 2013. New Academic Word list. <http://www.newgeneralservicelist.org/nawl-new-academic-word-list/> (01.08.2017).
- Bruce, Ian. 2011. *Theory and concepts of English for academic purposes*. NY: Palgrave Macmillian.

- Bruton, Anthony. 2007. "Vocabulary learning from dictionary reference in collaborative EFL translational writing". *System* 35, 353-367.
- Bussmann, Hadumod. 1996. *Dictionary of language and linguistics*. New York: Routledge.
- Campion Mary; Elley, Warwick. 1971. *An academic vocabulary list*. Wellington: New Zealand Council of Educational Research.
- Carroll, John; Davies, Peter; Richman, Barry. 1971. *The American Heritage Word Frequency Book*. New York: Houghton Mifflin.
- Carpenter, Shana; Olson, Kellie. 2010. "Are pictures good for learning new vocabulary in a foreign language? Only if you think they are not". *Journal of Experimental Psychology* 38(1), 92-101.
- Catalan, Jimenez. 2003. "Sex differences in L2 vocabulary learning strategies". *International Journal of Applied Linguistics* 13(1), 54-77.
- Chen Qi; Ge, Guang-chun. 2007. "A corpus-based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles (RA's)". *English for Specific Purposes* 26, 502-514.
- Cohen, Andrew D. 2011. *Strategies in learning and using a second language*. Harlow: Longman.
- 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.
- Cook, Linda; Mayer, Richard. 1983. "Reading strategies training for meaningful learning from prose". In Pressely, Michael; Levin, Joel. (eds.). *Cognitive Strategy Research*. New York: Springer, 87-131.
- Coxhead, Averil. 2000. "A New Academic Word List". *TESOL Quarterly* 34(2), 213-238.
- Coxhead, Averil. 2011. "The Academic Word List 10 Years On: Research and Teaching Implications". *TESOL Quarterly* 45(2), 355-362.
- Chung, Teresa; Nation, Paul. 2003. "Technical vocabulary in specialised texts". *Reading in a Foreign Language* 15(2), 103-116.
- Chung, Teresa; Nation, Paul. 2004. "Identifying technical vocabulary". *System* 32, 251-263.
- Curry, Lynn. 1983. "An organisation of learning styles theory and constructs". Paper presented at the Annual Meeting of the American Educational Research Association. Montreal Quebec, April 11-15), 2-25.
- Curry, Lynn. 1990. "A critique of the research on learning style". *Educational Leadership* 48(2), 50-54.
- De Chazal, Edward. 2014. *English for Academic Purposes*. Oxford: Oxford University Press.
- Demir, Ümmü. 2017. "The effects of pictures and sentence examples on foreign language vocabulary learning". *Journal of Foreign Language Education and Technology* 2(1), 24-38.
- Dibaj, Farzad .2011. "Vocabulary Learning; A comprehension of learners of English as a second language." *Australian Review of Applied Linguistics* 34(2), 193-215.
- Dörnyei, Zoltan. 2001. *Motivational strategies in the language classroom*. Cambridge: Cambridge University Press.
- Dörnyei, Zoltan. 2010. *Questionnaires in second language research: construction, administration, and processing*. New York: Routledge.
- Dunn, Rita; Dunn, Kenneth. 2005. "Thirty-Five Years of Research on Perceptual Strengths: Essential Strategies to Promote Learning". *Clearing House: A Journal of Educational Strategies, Issues and Ideas* 78(6), 273-276.
- Durrant, Philip. 2016. "To what extent is the Academic Vocabulary List relevant to university student writing?". *English for Specific Purposes* 43, 49-61.

- Durrant, Philip; Schmitt, Norbert. 2009. "To what extent do native and non-native writers make use of collocations?". *IRAL* 47(2), 157-177.
- Durrant, Philip; Schmitt, Norbert. 2010. "Adult learners' retention of collocations from exposure". *Second Language Research* 26(2), 163-188.
- Ellis, Nick C.; Beaton Alan. 1993. "Psycholinguistic determinants of foreign language vocabulary learning". *Language Learning* 43(4), 559-617.
- Ellis, Nick; Beaton, Alan. 1995. "Psycholinguistic determinants of foreign language vocabulary learning". In Harley, Birgit (ed.). *Lexical Issues in Language Learning*. Amsterdam: Benjamins, 107-165.
- Evert, Stefan. 2009. "Corpora and collocations". In Lüdeling, Anke; Kytö, Merja(eds.). *Corpus linguistics: an international handbook*. Berlin: Walter de Gruyter, 1212-1248.
- Francis, Nelson; Kucera, Henry. 1982. *Frequency Analysis of English Usage*. Boston MA: Houghton Mifflin.
- Gardner, Dee; Davies, Mark. 2014. "A new academic vocabulary list" *Applied Linguistics* 35(3), 305-327.
- Gathercole, Susan; Baddeley, Alan. 1989. "Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language* 28, 200-213.
- Ghadessy, Mohsen. 1979. "Frequency counts, word lists and material preparation: A new approach". *English Teaching Forum* 17(1), 24-27.
- Griffiths, Carol. 2008. "Strategies and good language learners." In Griffiths, Carol. (ed.). *Lessons from Good Language Learners*. Cambridge: Cambridge University Press, 83-98.
- Grinden, Ellen. 1992. *ANOVA: Repeated measures*. Newbury Park, CA: Sage.
- Gorjuan, Hayati; Barazandeh, Elham. 2012. "An evaluation of the effects of art on vocabulary learning through multi-sensory modalities". *Procedia Technology* 1, 345 - 350
- Gu, Yongqi. 2002. "Gender, academic major, and vocabulary learning strategies of Chinese EFL learners". *RELC Journal* 33(1), 35-54.
- Gu, Peter Yongqi. 2003. Vocabulary learning in a second language: person, task, context and strategies. *TESL-EJ* 7(2).
<http://www.tesle.org/wordpress/issues/volume7/ej26/ej26a4/> (20 April 2017).
- Gu, Peter Yongqi; Johnson, Robert. 1996. "Vocabulary Learning Strategies and Language Learning Outcomes". *Language Learning* 46(4), 643-679.
- Hirsh, David. 2010. *Academic Vocabulary in Context*. Bern: Peter Lang.
- Hong-Nam, Kyungsim; Leavell, Alexandra. 2006. "Language learning strategy use of ESL students in an intensive English learning context". *System* 34, 399-415.
- Hyland, Ken; Tse, Polly. 2007. "Is there an 'Academic vocabulary'?". *TESOL Quarterly* 41(2), 235-253.
- Ife, Anne; Vives Boix, Gemma; Meara, Paul. 2000. "The impact of study abroad on the vocabulary development of different proficiency groups". *Spanish Applied Linguistics* 4, 55-84.
- Jordan, Robert. 1997. *English for Academic Purposes: A guide and resource book for teachers*. Cambridge: Cambridge University Press.
- Keikhaie Yahya; Khoshkhoonejad, Amirali; Mansoorzadeh, Noorullah; Panahandeh, Essa. 2015. "Bilingualism and Vocabulary Learning: A Comparison between Baluchi and Persian EFL Learners". *Journal of Education and Learning* 9(3), 183-189.
- Kinginger, Celeste. 2011. "Enhancing Language Learning in Study Abroad". *Annual Review of Applied Linguistics* 31, 58-73.

- Kwary, Deny. 2011. "A hybrid method for determining technical vocabulary". *System* 39, 175-185.
- Llanes, Angles; Munoz, Carmen. 2009. "A short stay abroad: does it make a difference?" *System* 37(3), 353-365.
- Laufer, Batia; Hulstijn, Jan. 2001. "Incidental Vocabulary Learning in a second language: The construct of task-induced involvement". *Applied Linguistics* 22(1), 1-26.
- Lehecka, Tomas. 2015. "Collocation and colligation". In Jan-Ola, Östman; Jef, Verschueren (eds.). *Handbook of Pragmatics Online*. Amsterdam: Benjamins, 1-23.
- Word Smith Tools. 2017. (Version 7.0). [Computer Program]. Lexical Analysis Software Ltd; Oxford University Press <http://www.lexically.net/wordsmith/> (2. Aug. 2017).
- Lynn, Robert. 1973. "Preparing Word lists: A suggested method". *RELC Journal* 4(1), 25-32.
- Martin, Anne. 1976. "Teaching academic vocabulary to foreign graduate students". *TESOL Quarterly* 10(1), 91-97.
- Matthews, Peter H. 2014. *The concise Oxford Dictionary of linguistics*. Oxford: Oxford University Press. <http://www.oxfordreference.com.uaccess.univie.ac.at/view/10.1093/acref/9780199675128.001.0001/acref-9780199675128> (14. July 2017).
- Meara, Paul & Miralpeix, Imma. V_Size Manual. (14. July 2017).
- Meara, Paul & Miralpeix, Imma. D-Tools Manual. (14. July 2017).
- Meara, Paul & Miralpeix, Imma. V_Size. (Version v2.0) [Computer Program]. <http://www.lognostics.co.uk/tools/> (2. Aug. 2017).
- Meara, Paul & Miralpeix, Imma. 2015. D-Tools. (Version v2.00) [Computer Program]. <http://www.lognostics.co.uk/tools/> (2. Aug. 2017).
- Meyerhoff, Miriam; Schlee, Erik; MacKenzie, Laurel. 2015. *Doing Sociolinguistics. A practical guide to data collection and analysis*. New York: Routledge.
- Moon, Rosamund. 1997. "Vocabulary connections: multi-word items in English". In: Schmitt, Norbert; McCarthy, Michael (eds.). *Vocabulary: Description acquisition and pedagogy*. New York: Cambridge University Press, 40-63.
- Nagy, William. 1997. "On the role of context in first-and second-language learning". In Schmitt, Norbert; Mc Carthy, Michael (eds.). *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press, 64-83.
- Nakata, Tatsuya. 2008. "English vocabulary learning with word lists, word cards and computers: implications from cognitive psychology research for optimal spaced learning". *ReCALL* 20(1), 3-20.
- Nation Paul. 2005. *Range and Frequency. Programs for Windows based PCs*. [Computer Program] <http://www.victoria.ac.nz/lals/resources/range> (14. July 2017).
- Nation, Paul. 1990. *Teaching and Learning Vocabulary*. New York: Newbury House.
- Nation, Paul. 2013. *Learning Vocabulary in Another Language*. (2nd edition). Cambridge: Cambridge University Press.
- Nation, Paul; Gu, Peter Yongqi. 2007. *Focus on vocabulary*. Sydney: Macquarie University.
- Nation, Paul; Kyongho, Hwang. 1995. "Where would general service vocabulary stop and special purpose vocabulary begin?". *System* 23(1), 35-41.
- Nation, Paul; Waring, Robert. 1997. "Vocabulary size, text coverage and word lists". In: Schmitt, Norbert; McCarthy, Michael (eds.). Cambridge: Cambridge University Press, 6-19.
- Nyikos, Martha. 2008. "Gender and good language learners". In Griffiths, Carol (ed.). *Lessons from good language learners*. Cambridge: Cambridge University Press, 73-82.
- Ogawa, Daichi; Ikeno, Sakiko; Okazaki, Ryuta; Hachisu, Taku; Kajimoto, Hiroyuki. 2014. "Tactile cue presentation for vocabulary learning with keyboard". *UIST* 14, 5-8.

- Omid, Ahmad; Rashidi, Nasser. 2011. "A Survey on Iranian EFL Learners' Beliefs on the Role of Rote Memorization in Learning Vocabulary and Its Effect on Vocabulary Achievement". *Pan-Pacific Association of Applied Linguistics* 15(1), 139-161.
- Oxford English Online Dictionary. 2010. (3rd edition). Oxford: Oxford University Press. (14. July 2017).
- Oxford, Rebecca. 1990. *Language Learning Strategies: What every teacher should know*. Boston: Heinle & Heinle.
- Oxford, Rebecca; Crookall, David. 1990. "Vocabulary Learning: A Critical Analysis of Techniques". *TESL CANADA JOURNAL* 7(2), 9-30.
- Palmer, Harold. 1921. *The Principles of Language Study*. London. George G. Harrap & Co.
- Paquot, Magali. 2010. *Academic Vocabulary in Learner Writing: From Extraction to Analysis*. London: Continuum International Publishing Group.
- Praninskas, Jean. 1972. *American University Word List*. London: Longman.
- Purpura, James. 1999. *Learner Strategy use and performance on language tests: A structural equation modelling approach*. Cambridge: Cambridge University Press.
- Reid, Joy. 1987. "The learning style preferences of EFL students". *TESOL Quarterly* 21(1), 87-111.
- Ridely, Jennifer. 1997. *Reflection and strategies in foreign language learning*. Philadelphia: Multilingual Matters.
- Rubin, Joan. 1987. "Learner strategies: Theoretical assumptions, research history and typology". In Wenden, Anita.; Rubin Joan. (eds.). *Learner strategies in Language Learning*. New York: Prentice Hall, 15-30.
- Sanaoui, Razika. 1995. "Adult Learners' approaches to learning vocabulary in second languages". *The Modern Language Journal* 79(1), 15-28.
- Schmitt, Norbert. 1997. "Vocabulary learning strategies". In: Schmitt, Norbert; McCarthy, Michael (eds.). *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press, 199-227.
- Seong-Yeon. 2007. "Facilitating Vocabulary with pictures". *영어교육* 62(3), 27-53.
- Sinclair, John. 2004. "New evidence, new priorities, new attitudes". In: Sinclair, John (ed.). *How to use corpora in language teaching*. Amsterdam: John Benjamins, 271-299.
- SPSS. 2016. (Version 23). [Computer Program]. IBM.
- Stubbs, Michael. 1986. "Language development, lexical competence and nuclear vocabulary". In: Stubbs, Michael (ed.). *Educational Linguistics*. Oxford: Blackwell, 98-115.
- Syanova, Anna; Schmitt, Norbert. 2008. "L2 Learner Production and Processing of collocation: A multi-study". *The Canadian Modern Language Review* 64(3), 429-458.
- Takac, Visnja. 2008. *Vocabulary Learning Strategies and foreign language acquisition*. Frankfurt: Multilingual Matters Ltd.
- Tan, Po. 2011. "Towards a Culturally Sensitive and Deeper Understanding of "Rote Learning" and Memorisation of Adult Learners". *Journal of Studies in International Education* 15(2), 124-145.
- Thorndike, Edward. 1908. "Memory for paired associates". *Psychological Review* 15, 122-138.
- Thorndike, Edward; Lorge, Irving. 1944. *The Teacher's Word Book of 30,000 Words*. Columbia: Columbia University.
- Townsend, Dianna; Kiernan, Darl. 2015. "Selecting Academic Vocabulary Words Worth Learning". *The Reading Teacher* 69(1), 113-118.
- Unaldi, Ihsan; Bardakci, Mehmet; Akpinar, Kadriye Dilek; Dolas, Ferhat. 2013. "A comparison of contextualized, decontextualized and corpus-informed vocabulary

- instruction: A quasi-experimental study". *Journal of Language and Literature Education* 8, 78-95.
- Valipouri, Leila; Nassaji, Hossein. 2013. "A corpus-based study of academic vocabulary in chemistry research articles". *Journal of English for Academic Purposes* 12, 248-263.
- Van Wyk, Jacolynn; Mostert, Maria Louise. 2016. "The influence of mother tongue and gender on the acquisition of English (L2). The case of Afrikaans in Windhoek schools, Namibia". *Cogent Education*, 1-12.
- Vann, Roberta; Abraham, Roberta. 1990. "Strategies of unsuccessful language learners". *TESOL Quarterly* 24(2), 177-198.
- Weiß, Silvia; Tritischer-Archan, Sabine. 2014. "Das österreichische Bildungssystem [The Austrian Education System]". Insitut für Bildungsforschung der Wirtschaft, 1-4.
- In Schmitt, Norbert (ed.). *Formulaic Sequences*. Amsterdam: John Benjamins, 249-268.
- Webb, Stuart. 2007. "The effects of repetition on vocabulary knowledge". *Applied Linguistics* 28(1), 46-65.
- Wenden, Anita. 1987. "Conceptual background and utility". In Wenden Anita.; Rubin, Joan. (eds.). *Learner Strategies in Language Learning*. New York: Prentice Hall, 3-14.
- Wray, Alison. 2000. "Formulaic sequences in second language teaching: Principles and practice". *Applied Linguistics* 21(4), 463-489.
- Wray, Alison. 2002. *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Wray, Alison. 2004. "'Here is one I prepared earlier': Formulaic language learning on television".
- West, Michael. 1953. *A general service list of English words: with semantic frequencies and a supplementary word-list for the writing of popular science and technology*. London: Longman.
- British Council. *Word Family Framework*. [Computer Program].
<https://www.learnenglish.org.uk/wff/index.html>
- Xue, Guo-yi; Nation, Paul I. 1984. "A University Word list". *Language Learning and Communication* 3(2), 215-229.
- Yang, Weidong; Dai, Weiping. 2011. "Rote Memorization of Vocabulary and Vocabulary Development". *English Language Teaching* 4(4), 61-64.
- Zare, Mostafa; Mobarakeh, Davoudi. 2013. „Effects of Bilingualism on L3 Vocabulary Learning among Iranian EFL Learners". *GEMA Online Journal of Language Studies* 13(1), 127-128.

6. APPENDIX

6.1. Abstract

This study is specifically concerned with the relationship between academic, biographical background of participants and vocabulary learning strategies applied.

The participants of this study are thirty-seven students of the English Department of the University of Vienna, who provided three texts from courses in the language competence program at different stages of their studies. These texts were further analysed with the help of the corpus analysis program AntProfiler to gather information on the academic, high-frequency, technical and low-frequency vocabulary percentages in the texts. In addition, an online questionnaire was used to collect data on vocabulary learning strategies used by the participants.

The results clearly indicated that the participants were able to increase their academic vocabulary, technical and low-frequency vocabulary throughout their studies, while high-frequency vocabulary decreased indirect proportionally. However, different levels of proficiency in academic vocabulary use could be detected. Concerning the relationship between biographical background information provided, such as gender, bilingualism or experience abroad, and vocabulary development no statistically significant results could be detected. The same accounts for the relationship between vocabulary learning strategies applied and academic vocabulary development. However, it was possible to find correlations between vocabulary learning strategies and biographical background information on gender, existence of English relatives and school type attended. A qualitative analysis revealed that accuracy of vocabulary use and formality increased from the beginning to the end of the studies.

These findings suggest that the students of the English Department of the University of Vienna manage to become more competent participants in the academic discourse community. Moreover, the results imply that academic vocabulary development is not linked to vocabulary learning strategies used or biographical background information, which might stimulate controversy in the field and needs to be replicated and examined more closely in comparable studies.

6.2. Zusammenfassung

Diese Studie beschäftigt sich mit einem möglichen Zusammenhang zwischen der Entwicklung des akademischen und hochfrequenten Vokabulars am Englisch Department der Universität Wien und den biographischen Hintergründen der Teilnehmer, sowie mit den Vokabellernstrategien, die angewendet wurden.

Siebenunddreißig Englisch Studenten der Universität Wien wurden in diese Studie miteingebunden. Jeder Teilnehmer stellte drei englische Texte zur Verfügung, die in drei verschiedenen Kursen des Sprachkompetenz Programms im Laufe des gesamten Studiums verfasst wurden. Diese Aufsätze wurden anschließend mit Hilfe des Corpus Analyse Programmes AntProfiler analysiert, um Information über die prozentuelle Verteilung von akademischen und hochfrequenten Vokabeln zu erhalten. Des Weiteren, wurden Daten zu den von den Teilnehmern verwendeten Vokabellernstrategien mit Hilfe eines online Fragebogens gesammelt.

Die Resultate zeigen klar, dass die Studenten ihr akademisches, technisches und niedrig frequentes Vokabular im Laufe ihres Studiums statistisch signifikant steigern konnten, wozu sich das hochfrequente Vokabular indirekt proportional verhielt. Bei genauerer Analyse konnten jedoch unterschiedliche Kompetenzniveaus hinsichtlich der Vokabelverwendung festgestellt werden. Bezüglich des Zusammenhanges zwischen Vokabelverwendung und biographischen Hintergrundinformationen konnten keine signifikanten Korrelationen gefunden werden. Dasselbe galt für eine mögliche Verbindung zwischen Vokabelverwendung und den angewendeten Vokabellernstrategien. Die Ergebnisse bestätigten jedoch einen statistisch relevanten Zusammenhang zwischen den biographischen Angaben hinsichtlich Gender, Existenz von englischsprachigen Verwandten und besuchtem Schultypus und bestimmten Vokabellernstrategien. Eine qualitative Analyse zeigte des Weiteren, dass die Formalität der gewählten Wörter im Laufe des Studiums anstiegen und die Vokabelfehler reduziert wurden.

Diese Resultate zeigen, dass es den Studenten am Englisch Department der Universität Wien gelingt, sich zu kompetenteren Mitgliedern der akademischen Gesellschaft heranzubilden. Des Weiteren, liefert die Studie kontroverse Ergebnisse hinsichtlich der mangelnden Verbindung zwischen Vokabular Entwicklung und Vokabellernstrategien, die in weiterfolgenden Studien genauer untersucht werden sollte.

Academic Vocabulary Proficiency, Vocabulary Learning Strategies and Biographical Background

Dear Colleagues,

Thank you very much for participating in my study. As you know, this questionnaire is one part of my diploma thesis about academic vocabulary. The second part is three uncorrected texts (1 opinion essay from ILSS1, 1 opinion essay from ILSS2 and the book review from EAP) that you send me via e-mail. To match your texts with your questionnaire as anonymously as possible I will ask you to indicate your matriculation number at the beginning of the questionnaire and on your texts. Therefore, even I won't know whose texts or questionnaire I am analysing and your data will be handled strictly anonymously and confidentially. The following questionnaire is no test. There are no right or wrong answers, so try to answer as honestly as possible. Thank you very much in advance for your support and time.

Your matriculation number: _____

Part 1: Vocabulary learning strategies

a) Planning vocabulary learning

The following section describes planning strategies that can be used for vocabulary learning. Read each statement and mark the option that best suits you, according to the following scale:

- 1 means you always use the learning strategy described
- 2 means you almost always use it
- 3 means you often use it
- 4 means you sometimes use it
- 5 means you rarely use it
- 6 means you never use it

1. I consciously choose the words I want to learn.
1 2 3 4 5 6
2. I consider the aspects of a word which are most useful for me.
1 2 3 4 5 6
3. I have a big repertoire of vocabulary learning strategies.
1 2 3 4 5 6
4. I know how to choose, adapt and combine my vocabulary strategies.
1 2 3 4 5 6
5. I plan revisions of vocabulary items regularly and I have a strategy for doing so.
1 2 3 4 5 6

b) Recording vocabulary

The following section describes recording strategies that can be used for learning vocabulary. Read each statement and mark the option that best suits you, according to the following scale:

- 1 means you always use the learning strategy described
- 2 means you almost always use it
- 3 means you often use it
- 4 means you sometimes use it
- 5 means you rarely use it
- 6 means you never use it

1. I do not write down vocabulary in an organised form. I just repeat items orally.
1 2 3 4 5 6
2. I do not collect vocabulary in an organised form. I repeat it by writing it down.

- | | | | | | | |
|----|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 3. | I keep a word list. | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 4. | I make flash cards. | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 5. | I like to stick sheets with vocabulary written on them on objects in my surroundings. | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 6. | I keep a notebook or vocabulary log. | | | | | |
| | 1 | 2 | 3 | 4 | 5 | |
| 7. | Other strategy/ies that I use for recording: _____ | | | | | |

c) Discovering the meaning of an unknown word:

The following section describes strategies that can be used to discover the meaning of an unknown word. Read each statement and mark the option that best suits you, according to the following scale:

- 1 means you always use the learning strategy described
- 2 means you almost always use it
- 3 means you often use it
- 4 means you sometimes use it
- 5 means you rarely use it
- 6 means you never use it

If I do not know a word I help myself by...

- | | | | | | | |
|----|--|---|---|---|---|---|
| 1. | Guessing it's meaning from the context. | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| | a) Choose a suitable option/s | | | | | |
| | When guessing from context I look at | | | | | |
| | <ul style="list-style-type: none"> • the part of speech of the word • the roots and affixes of the word • the content of the surrounding text | | | | | |
| | b) I check if my guess was correct. | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 2. | Try to recall a similar word in my L1 | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 3. | Try to recall a similar word in other languages I know. | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 4. | Analysing available illustrations. | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 5. | Consulting a reference source. | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| | When consulting a reference source, I use: | | | | | |
| | <ul style="list-style-type: none"> • a monolingual dictionary | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| | <ul style="list-style-type: none"> • a bilingual dictionary | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| | <ul style="list-style-type: none"> • a language corpus | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| | <ul style="list-style-type: none"> • an existing word list. | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 6. | Simply ignoring or skipping the word. | | | | | |

- | | | | | | | |
|-----|--|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 7. | Looking for any defining words or phrases in the rest of the text. | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 8. | Asking a teacher. | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 9. | Asking my classmates or friends. | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 10. | Asking a native speaker I know. | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |

d) Learning vocabulary:

The following section describes vocabulary learning strategies that can be used. Read each statement and mark the option that best suits you, according to the following scale:

- 1 means you always use the learning strategy described
 - 2 means you almost always use it
 - 3 means you often use it
 - 4 means you sometimes use it
 - 5 means you rarely use it
 - 6 means you never use it
 - 7 means you have never heard of it.
- | | | | | | | | |
|-----|--|---|---|---|---|---|---|
| 1. | I learn vocabulary together with other learners of English. | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2. | I learn vocabulary with a native speaker. | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 3. | I act out role plays with other learners to practice vocabulary. | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 4. | I prepare presentations or speeches incorporating vocabulary items that I want to learn. | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 5. | I learn words in combination with any kind of picture or visual representation. | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 6. | When I learn a word I visually imagine the meaning of the word. | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 7. | I connect the word with a personal experience to remember it. | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 8. | I use semantic mapping systems, such as a mind map or clustering. | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 9. | I learn a word together with its synonyms. | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 10. | I learn a word together with its antonyms. | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 11. | I learn any irregular inflectional morphology at the same time. | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 12. | I use the PEG Method | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 13. | I use the Keyword Method. | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 14. | I group words spatially to remember them. | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 15. | I use a new word in a sentence. | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 16. | I embed new words in a storyline to remember them. | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 17. | I learn spelling specifically. | | | | | | |

- 1 2 3 4 5 6 7
18. I learn the pronunciation of a word by speaking it out loud.
1 2 3 4 5 6 7
19. I learn the phonetic transcription.
1 2 3 4 5 6 7
20. I imagine the graphical form of the word. e.g.: elephant
1 2 3 4 5 6 7
21. I underline the initial letter of a word.
1 2 3 4 5 6 7
22. I try to remember the word in the context in which I encountered it.
1 2 3 4 5 6 7
23. I remember individual words which I can combine to larger chunks.
1 2 3 4 5 6 7
24. I learn cognates/translations.
1 2 3 4 5 6 7
25. I learn collocations and colligations for a new word.
1 2 3 4 5 6 7
26. I use physical action to remember a word.
1 2 3 4 5 6 7
27. I use semantic feature grids.
1 2 3 4 5 6 7
28. I use pre-existing material to practice vocabulary.
1 2 3 4 5 6 7
29. I paraphrase the word I want to learn.
1 2 3 4 5 6 7
30. I intentionally interact with a native speaker/friend in English to increase my vocabulary knowledge.
1 2 3 4 5 6 7
31. I intentionally translate language sequences in written or spoken form into my L1/L2 to increase my vocabulary knowledge.
1 2 3 4 5 6 7
32. I intentionally use extensive reading to increase my vocabulary knowledge.
1 2 3 4 5 6 7
33. I intentionally use glossing to increase my vocabulary knowledge.
1 2 3 4 5 6 7
34. I intentionally listen to English.
1 2 3 4 5 6 7
35. I intentionally watch movies to increase my vocabulary knowledge.
1 2 3 4 5 6 7
36. I intentionally write texts in English to increase my vocabulary knowledge.
1 2 3 4 5 6 7
37. Any other vocabulary learning strategy? _____

e) Self-management strategies:

The following section describes self-management strategies that can be used for vocabulary learning. Read each statement and mark the option that best suits you, according to the following scale:

- 1 means you always use the learning strategy described
- 2 means you almost always use it
- 3 means you often use it
- 4 means you sometimes use it
- 5 means you rarely use it
- 6 means you never use it

1. I regularly remind myself of my vocabulary learning goals

1 2 3 4 5 6

Name strategies you use to remind yourself of goals: _____

2. I regularly remind myself of the value of achieving success in vocabulary learning.

1 2 3 4 5 6

Name strategies you use to remind yourself of the value of vocabulary learning: _____

3. I regularly monitor my concentration so as to learn vocabulary effectively.

1 2 3 4 5 6

Name strategies you use to monitor or regulate your concentration: _____

4. I adapt tasks to avoid boredom when learning vocabulary.

1 2 3 4 5 6

Name strategies you use to adapt tasks: _____

5. I use relaxation techniques to help me learn vocabulary.

1 2 3 4 5 6

Name strategies you use to relax: _____

6. I try to generate positive emotional reactions when learning vocabulary.

1 2 3 4 5 6

Name strategies you use to generate positive emotions: _____

7. I use self-encouragement techniques when learning vocabulary.

1 2 3 4 5 6

Name strategies you use to encourage yourself: _____

8. I adapt my environment to reach my goals when learning vocabulary.

1 2 3 4 5 6

How do you adapt your environment to reach your goals? _____

Part 2: Personal information

Choose the correct answer/s.

1. Gender: Male Female
2. Secondary school type attended: Neue Mittelschule/Hauptschule
Gymnasium Unterstufe
Gymnasium Oberstufe
HAK
HTL
HLW
HASCH
BAKIP
Abendschule
Externistenmatura
Other: _____

Finally please answer the following questions:

3. Have you spent a longer period of time abroad? If so: Where? When? For how long (indicate in months)? E.g.: London 2015, nine months.

4. What is your mother tongue? _____
5. Are you bilingual? If so: Which languages do you speak? _____
6. Do you have English speaking relatives? If yes: How often do you have the opportunity to speak with them? _____

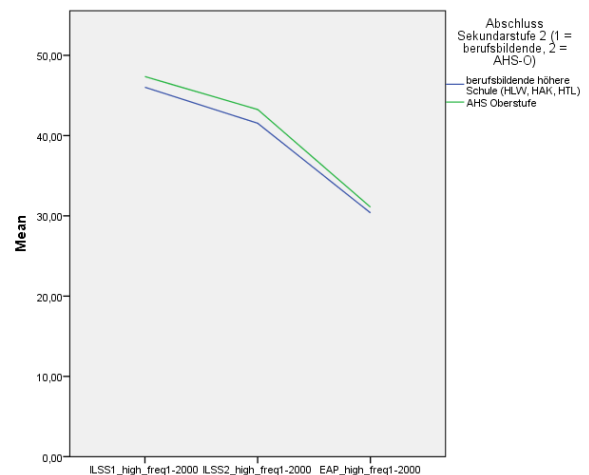
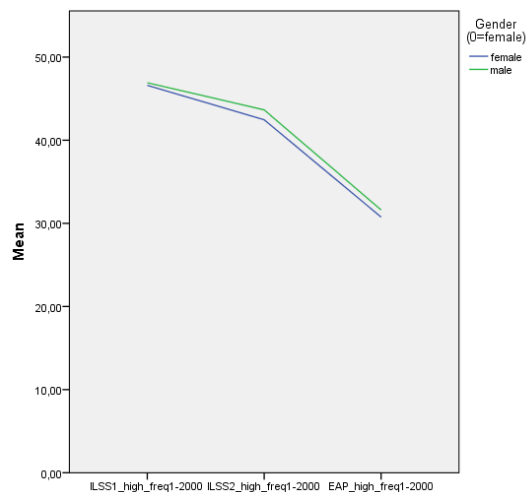
I tried to make this questionnaire as comprehensible as possible. However, if anything was left open that you urgently want to say you have the chance now:

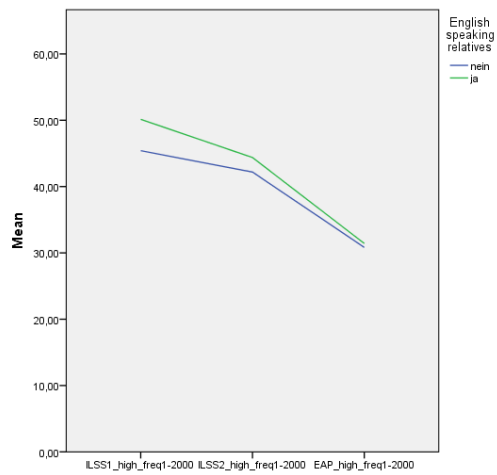
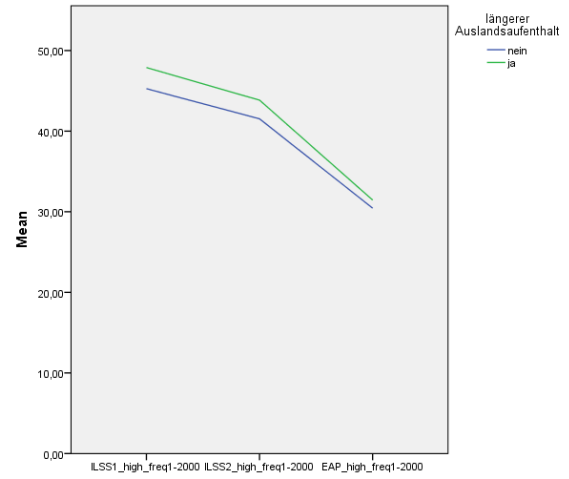
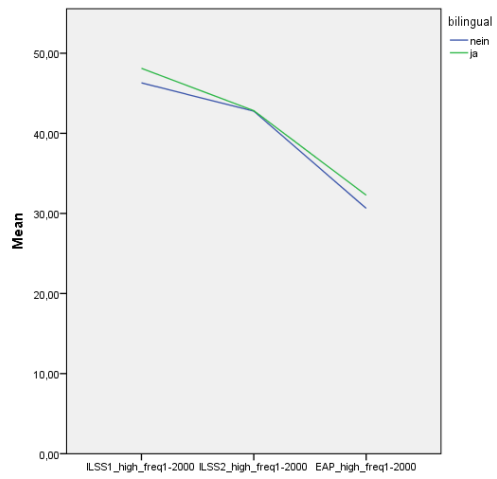
Thank you very much for your participation. If you want to be informed about my results in general or about your specific score and correlations, do not hesitate to write me an e-mail to K.ghamarian@gmx.at

PLEASE DO NOT FORGET TO SEND ME YOUR THREE TEXTS INCLUDING YOUR MATRICULATION NUMBER VIA E-MAIL AS WELL! Thank you very much.

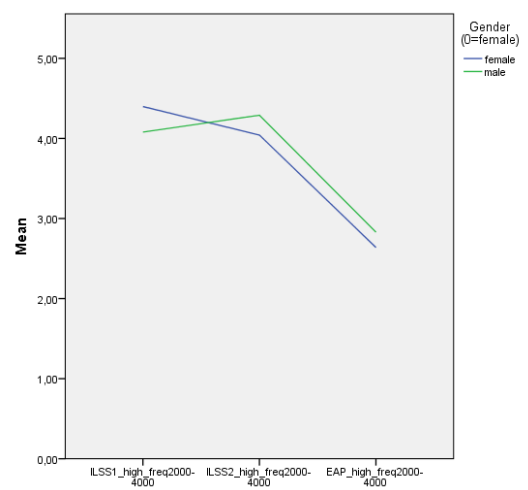
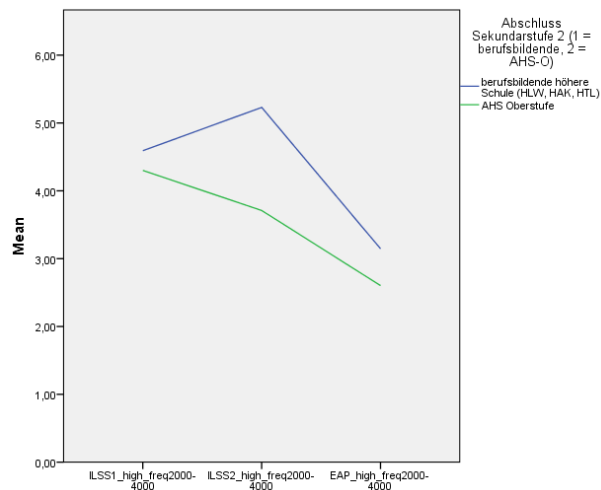
Graphs for correlations between vocabulary types and biographical data:

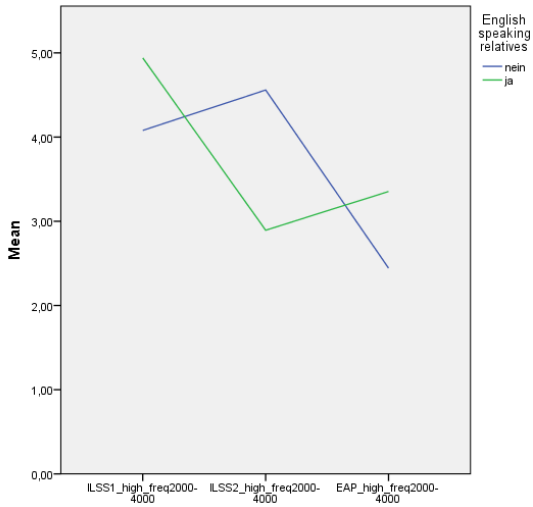
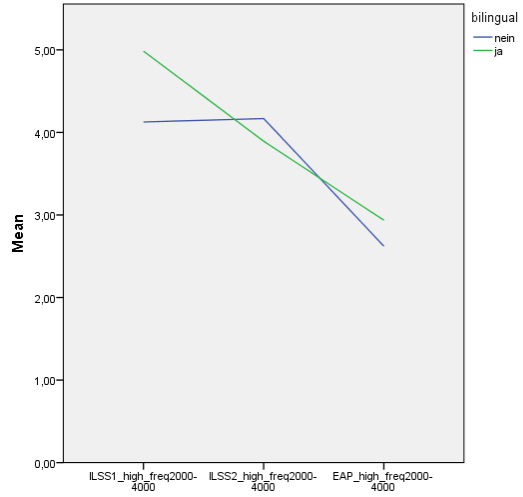
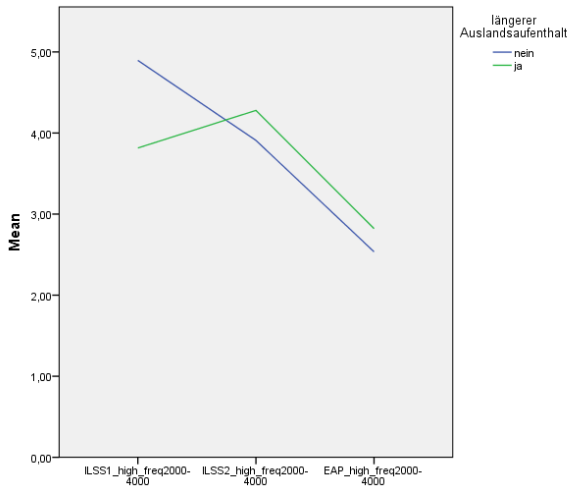
1. High-frequency (1-2000)





2. High-frequency vocabulary (2000-4000):





3. Academic vocabulary:

