



universität
wien

MASTERARBEIT

Titel der Masterarbeit

„Collocation Acquisition by Ukrainian/Russian L1
Learners of EFL “

Verfasserin

Maryna Dudnyk

angestrebter akademischer Grad

Master of Arts (MA)

Wien, 2013

Studienkennzahl lt. Studienblatt:

A 066 812

Studienrichtung lt. Studienblatt:

English Language and Linguistics

Betreuerin:

Univ.-Prof. Dr. Evelien Keizer

Table of contents

Abstract.....	1
1. Introduction.....	2
1.1 Setting the scene and statement of the problem.....	2
1.2 Aims and scope of the study	3
1.3 Research questions and hypotheses	6
1.4 Definition of terms.....	8
1.5 Thesis outline	9
2. Theoretical background and previous research.....	11
2.1 The phenomenon of collocation.....	11
2.1.1 Theoretical frameworks of collocation research.....	11
2.1.2 Definition of collocation in the study	13
2.1.3 Collocation knowledge: defining the construct	16
2.2 The importance of collocations in language learning.....	19
2.3 L2 knowledge and use of collocations.....	22
2.3.1 Corpus-based research	23
2.3.2 Elicitation studies.....	27
2.3.3 Summary of the reviewed studies and key findings	32
2.4 L2 collocation acquisition.....	35
2.4.1 Placing into a broader framework.....	35
2.4.2 Factors effecting acquisition: interpreting previous findings	39
2.4.3 Learning in the EFL classroom.....	43
2.5 Chapter summary	45
3. Research methodology.....	47
3.1 Setting	47
3.2 Participants.....	48
3.3 Test instrument and procedure.....	49
3.3.1 Test format	49
3.3.2 Material.....	50
3.3.3 Test administration.....	54
3.3.4 Data analysis and scoring procedure.....	55
3.4 Research hypotheses	56
4. Results and discussion	57
4.1 EFL learners' productive collocation knowledge.....	57

4.2 Degrees in productive knowledge of collocations	62
4.3 Frequency of collocations and their mastery	63
4.4 The quantity and quality of EFL input in learning collocations	65
4.4.1 Length of EFL exposure	67
4.4.2 Intensity of EFL exposure.....	68
4.4.3 The EFL classroom.....	70
5. Conclusions.....	74
5.1 Significance of the study.....	74
5.2 Overview of the main findings and conclusions.....	75
5.3 Pedagogical implications	77
5.4 Limitations of the study	80
5.5 Suggestions for further research	81
References.....	83
Appendix I	89
Appendix II.....	93
Abstract in German.....	101
Curriculum Vitae	103

Abstract

Although the significance of collocations in language learning as well as for fluent and natural-sounding language use is widely recognized, research has consistently shown that L2 learners experience considerable difficulty with target language collocations, particularly in producing them correctly and with the native speaker diversity (e.g. Bahns & Eldaw 1993; Bonk 2001; Gitsaki 1999; Granger 1998; Howarth 1996; Nesselhauf 2005). Despite ample evidence of the difficulty L2 collocations cause, however, the mechanisms of how non-native speakers acquire collocations, including factors that influence this process, remain an open question until now. Evidence exists that a significant factor in collocation acquisition can be the environment in which learners come to acquire the target language (e.g. Aghbar 1990; Alskaran 2011; Siyanova & Schmitt 2008; Yamashita & Jian 2010). This area, however, is still largely under-explored, with a conspicuous gap in research on collocation learning as it occurs in the English as a foreign language (EFL) environment, especially by native speakers of Russian and Ukrainian. The present thesis constitutes an attempt at filling this gap by investigating collocation acquisition by Ukrainian/Russian speaking advanced learners of EFL. The focus is on the analysis of the productive knowledge of non-congruent verb-noun collocations involving common verbs in delexical uses.

Two main, interconnected aims guided this study. The first aim was to evaluate EFL learners' productive collocation knowledge and its development from the first to the last academic year. The second aim was to explore some factors that might influence collocation learning in the EFL setting, namely the frequency of collocations, the quantity (length) and the quality (intensity and type) of exposure to L2 input. To address these aims a special test instrument was developed consisting of a 20-item cloze test and a 20-item translation task. The results of the analysis showed that non-congruent delexical verb-noun collocations pose a serious challenge to Ukrainian/Russian speaking learners of EFL even at an advanced level, whose implicit acquisition in the EFL classroom, although occurs, is very slow and unpredictable. Factors found to positively influence acquisition are the frequency of the items, and the quality of L2 input, namely extra-curricular EFL engagement and the type of EFL instruction, with the overriding importance of the latter variable. Taken together, the findings of the study strongly suggest that in order for high-learning burden collocations to be effectively mastered in the EFL classroom, rich and extensive exposure to L2 (implicit techniques) needs to be supplemented by special pedagogical treatment (explicit techniques).

1. Introduction

1.1 Setting the scene and statement of the problem

Despite increasing globalization, for many people worldwide the opportunities for learning English remain largely limited to the EFL environment. Ukraine, the researcher's homeland, is a case in point. At present, English is a compulsory foreign language subject at most Ukrainian schools and universities, and people learn the language for various reasons, e.g. to increase employment opportunities, study or live in a foreign country, or just for travelling. Regardless of the exact purpose, all EFL learners share a striving for communicative efficiency in English that is well-attested to heavily depend on the phrasal competence, namely the knowledge of second language (L2) collocations (Bahns & Eldaw 1993; Howarth 1996; Pawley & Syder 1983; Sinclair 1991). Collocations are also widely recognized to play an important role in language learning, enhancing both receptive and productive linguistic competence, as well as to contribute to native-like fluency and idiomaticity in L2 (Lewis 1993; Nation 2001; Nattinger & DeCarrico 1992; Wray 2002).

A good command of target language collocations can thus be regarded as an indispensable skill for L2 learners, especially if they are to become competent English translators/interpreters or teachers, as in the case of the participants of the present study. Despite the acknowledged benefits of learning collocations, this area has also been consistently found to pose a serious challenge to the L2 learner, particularly in the EFL context (e.g. Bahns & Eldaw 1993; Biskup 1992; Bonk 2001; Gitsaki 1999; Howarth 1996, 1998a; Laufer & Waldman 2011; Nesselhauf 2005). Research shows that learner language tends to be characterized by a significant underuse of L2 collocations and, on the other hand, an overuse of a very limited set of favored combinations, which makes learner language less idiomatic and versatile than that of English native speakers (e.g. Durrant & Schmitt 2009; Hasselgren 1994; Howarth 1996, 1998a; Kaszubski 2000).

In particular, learners' first language (L1) was found to strongly influence L2 collocation perception and production, often resulting in deviant/erroneous L2 combinations that do not decrease even at advanced levels or as a function of growth in L2 proficiency (e.g. Howarth 1996; Laufer & Waldman 2011; Nesselhauf 2005; Yamashita & Jian 2010). Particularly problematic for non-natives were proved to be verb-noun collocations that are restricted and/or L1-L2 non-equivalent, as well as those involving a common delexical verb (Altenberg & Granger 2001; Chi et al. 1994; Eftekhari & Rahimi 2011; Kaszubski 2000;

Nesselhauf 2005; Yamashita & Jian 2011). There is also empirical evidence suggesting that the linguistic environment in which non-natives acquire the target language might have either a positive or an adverse impact on collocation acquisition (Alsakran 2010; Nesselhauf 2005; Siyanova & Schmitt 2008; Yamashita & Jian 2011).

On the basis of their findings, researchers conclude that non-natives' poor collocation knowledge may be first of all due to a lack of awareness of the syntagmatic relationships in L2. This, in turn, can be explained by the nature of collocations. Due to their semantic transparency, collocations are not usually attended to when met in the input, which is, however, an essential requirement for building associations between items that customarily co-occur. In other words, meaningful associations between L2 words encountered in the input usually remain unnoticed by the learner and hence are not amenable to chunking learning. It is therefore generally agreed that collocations cannot be left to chance in the classroom, in the hope they will be picked up merely from exposure to L2 input, but necessarily require pedagogical treatment (e.g. Al-Zahrani 1998; Bahns & Eldaw 1993; Eftekhari & Rahimi 2011; Farghal & Obiedant 1995; Nesselhauf 2005; Schmitt 2010; Yamashita & Jian 2011).

Despite numerous valuable insights into L2 learners' collocation knowledge and use, research in the field vary widely in terms of methods, types of target collocations, as well as participants' L1 and L2 proficiency profiles and learning environments. Due to this, it is often very difficult to compare the findings and draw valid conclusions. Moreover, the area of collocation acquisition taking place in the EFL classroom remains largely under-explored. Based on the author's learning/teaching experience, it can be speculated that collocations receive only little attention in the EFL curriculum and teaching materials in Ukraine. Apart from this impressionistic observation, however, there are a large number of open questions to be addressed, which would help make Ukrainian EFL classrooms more effective environments in terms of learning and teaching collocations. The present investigation can thus be considered a valuable contribution to the field by exploring collocation acquisition by Ukrainian/Russian speaking learners of EFL, including the analysis of the influence of two different EFL learning settings, i.e. two programs of study, (Translation/Interpreting and Teacher Training) in two institutions located in two Ukrainian cities, Cherkassy and Kiev.

1.2 Aims and scope of the study

The present study intends to investigate the acquisition of verb-noun collocations by advanced adult Ukrainian/Russian L1 learners in the EFL context, using two elicitation techniques, a cloze test and a translation task. More specifically, the study has two main,

interconnected aims. The first aim is to evaluate learners' productive collocation knowledge (cued and free active) and its development from the first to the last academic year. The second aim is to explore some factors that might influence collocation learning, in particular the frequency of collocations, the quantity (length) and quality (intensity and type) of exposure to EFL input. A secondary aim is to formulate suggestions for teaching collocations based on the results obtained. Considering that L2 collocation research is relatively new to Ukraine, the study intends to be largely exploratory in nature in that it investigates L2 collocation knowledge in the Ukrainian higher education system with the aim to reveal some key issues on the subject to be addressed with more rigor in further research, as well as to draw Ukrainian EFL educators' attention to the significance of collocations in learning and teaching. The investigation can thus be viewed as a study in the field of Applied Linguistics as it is aimed to be applied in or to inform English language teaching in the (Ukrainian) EFL classroom.

The study is primarily focused on investigating external factors in EFL acquisition of collocations, which means that such learner-related variables as motivation, language aptitude, or socio-cultural aspects (e.g. learner's age, gender or social status) are largely beyond the scope of the investigation. The learner sample in the study is restricted to Ukrainian/Russian L1 university level learners of EFL. Furthermore, the analysis is limited to the productive knowledge of incongruent verb-noun collocations involving common verbs in delexicalized uses. This type of collocation has been found to be particularly difficult for L2 learners, as well as to induce considerable L1 transfer, resulting in the production of deviant combinations. The focus on the delexicalized verb element, in particular, is motivated by evidence from the COBUILD project showing that most common English words tend to co-occur in delexical meanings (Barfield & Gyllstad 2009; Sinclair 1991:113), which makes such verbs especially important for the L2 learner.

The verb-noun combination type was chosen on several grounds. First of all, such combinations are not only frequent (Bahns & Eldaw 1993; Howarth 1996) and also among the most difficult ones for learners (Biskup 1992; Howarth 1996, 1998a; Nesselhauf 2005), but also indispensable in discourse as they carry core communicative information in utterances (Altenberg 1993: 227). With verb-noun combinations it is also often not possible for the learner to omit or paraphrase the difficult element or replace it with a safe choice (such as *very*) as in the case of adverb-adjective or adjective-noun collocations (e.g. Bahns & Eldaw 1993; Granger 1998). In general, the advantage of focusing on collocations in researching the L2 use of phraseology is that findings about this type of multi-words are applicable to a wider area of formulaic language (Barfield & Gyllstad 2009: 237).

Furthermore, the focus is on the productive knowledge aspect because it is generally assumed to be more challenging than comprehension even for advanced learners (Laufer 1998). The incremental nature of vocabulary/collocation learning implies that also productive knowledge of collocations develops gradually, that is from the passive, as in prompted production, to more active, independent and spontaneous use of collocations. Empirical support for this pattern can be found in Farghal & Obiedat's (1995) study showing non-natives to perform conspicuously better on a cloze test (cued production) than on translation (free production) that was obviously more difficult. Thus, to enable a fine-grained analysis of the development of the productive aspect of L2 collocation knowledge, the present investigation employed two productive collocation measures: a cloze test and a translation test to tap the cued and the free productive knowledge of collocations, respectively.

Restricting the analysis to a single L1 group was deemed necessary since evidence exists that the first language plays a significant role in perceiving and learning L2 collocations (Biskup 1992; Hasselgren 1994; Nesselhauf 2005; Yamashita & Jian 2011). Moreover, given that Russian/Ukrainian has rarely been examined in previous research, investigating L2 learners with Russian/Ukrainian L1 can be regarded as a justified and necessary contribution to the developing field, furthering understanding about L2 collocation acquisition as affected by languages different from those already explored. Finally, as mentioned earlier, the study is conducted within the context of foreign language learning, also investigating two different types of EFL classrooms: translation/interpreting and English teacher training programs in two institutions.

To the best of my knowledge, there are no studies so far that differentiated between L2 learners depending from what EFL learning setting (program of study) they stemmed. In most cases, EFL learners with varying L2 learning profiles, as well as L1s, were assumed to either represent the same learner population or were distinguished from the ESL or the native-speaker learner group. Evidence exists, however, that the learning environment can be a significant variable in collocation acquisition as L2 learning settings (ESL and EFL) can vary greatly in terms of the quantity and quality of L2 input, including methodologies and approaches to EFL teaching (Gass & Selinker 2008; Laufer & Girsai 2008; Wray 2002). Following this, it is logical to assume that differences exist also between classrooms within the single EFL learning context, e.g. between different EFL curricular, which is explored in the present study.

The idea to investigate the influence of two different EFL micro-contexts on learning collocations originally comes from the researcher's personal interest in testing a widely held belief in Ukraine that some universities are better than others in terms of teaching English,

and hence in preparing future specialist in the target language. Since a good command of L2 collocations is viewed as a hallmark of high proficiency in L2, it was assumed that assessing collocational competence of learners studying at a given university can shed some light on the institution's efficiency as regards EFL teaching.

1.3 Research questions and hypotheses

The two main aims outlined above are operationalized into overall six research questions (RQs) including the hypotheses (Hs). The first (RQ1) and the second research questions (RQ2) relate to learners' general and nuanced (cued and free) productive knowledge of collocations at two different levels of L2 competence (in years 1 and 4).

RQ1: Is there a significant difference between the first-year and the last-year EFL students in their productive collocation knowledge?

Research has consistently found a positive relationship between an increase in learners' general L2 proficiency and collocation knowledge, so that more advanced learners demonstrate an overall better collocation mastery (e.g. Bonk 2001; Deogratias 2011; Fatemeh & Arshad 2011; Gitsaki 1999; Zhang 1993). In light of this finding, it can be assumed that the Ukrainian/Russian EFL students will show an observable increase in productive collocation knowledge over the four-year period of formal instruction.

H1: Last-year EFL students demonstrate distinctly better productive collocation knowledge than students in their first academic year.

RQ2: Is there a significant difference between participants' collocation knowledge at the levels of cued production and free production?

The incremental nature of vocabulary and collocation learning (Li & Schmitt 2009; Schmitt 2010) together with evidence suggesting that free production is more difficult than cued production or comprehension (e.g. Alsakran 2011; Farghal & Obiedat 1995; Laufer 1998) provide a good reason for hypothesizing that the participants will better perform on the cloze test than on the translation task.

H2: Participants' knowledge of collocations at the free production level lags behind their cued productive collocation competence, which is more difficult to achieve.

The third research question (RQ3) relates to mastery of collocations depending on their frequency.

RQ3: Is there a significant difference between the subjects' performance on high-frequency and on low frequency collocations?

Positive effects of the frequency of collocations on their acquisition and perception are widely attested in the literature (e.g. Deogratias 2011; Durrant 2008; Durrant & Schmitt 2009; Siyanova & Schmitt 2008), based on which it is expected that more frequent collocations are better acquired than less common combinations, causing more difficulties.

H3: Learners' performance on high-frequent collocations is better than on low-frequent combinations that are more difficult to master.

Research questions 4 to 6 concern the role of the quantity and quality of EFL input in learning collocations. The fourth research question (RQ4) addresses the factor of the quantity (length) of exposure to EFL.

RQ4: Is there a relationship between the participants' knowledge of collocations and the length of exposure to EFL?

The importance of extensive exposure to L2 input for effective language learning is well documented (Coady & Huckin 1997; Durrant & Schmitt 2009, 2010; Nation 2001; Siyanova & Schmitt 2008). At the same time, there is evidence suggesting that, particularly in the EFL context as well as with high-learning burden collocations, considerably longer L2 exposure might be necessary to yield stable and significant progress, because even advanced students tend to experience difficulty with such collocations (Al-Zahrani 1998; Laufer & Waldman 2011; Nesselhauf 2005; Yamashita & Jian 2010). Moreover, numerous studies have reported that the use of collocations continues to cause difficulty to learners irrespective of the amount of L2 input they receive (Bahns & Eldaw 1993; Gitsaki 1999; Eftekhari & Rahimi 2011; Howarth 1996; Laufer & Waldmann 2011). It is therefore reasonable to expect that the 4-year difference in EFL classroom exposure will not substantially distinguish between the learners' better and worse collocation performance.

H4: There is no (or a very weak) relationship between the participants' length of EFL exposure and their command of collocations.

The fifth research question (RQ5) explores the effect of the quality (intensity and type) of exposure to EFL.

RQ5: Is there a relationship between the participants' production of collocations and their extra-curricular exposure to EFL?

It is generally recognized that high-quality input as well as intensive engagement with L2 leads to better learning gains (e.g. Li & Schmitt 2009; Schmitt 2010). It is thus reasonable to expect that students who learn and practice the target language in addition to classroom instruction, will normally demonstrate better collocation knowledge than those with ordinary EFL exposure.

H5: Participants with extra-curricular EFL engagement show better collocation performance than those whose L2 exposure is limited to the classroom EFL input.

The sixth research question (RQ6) addresses the quality (intensity and richness) and the type of exposure to EFL input, where the type refers to a particular EFL learning environment, in this case a study program, which differs from other settings in the emphasis of its curriculum.

RQ6: Is there a difference in the production of collocations between the subjects from different EFL learning settings?

Contexts in which non-natives come to acquire an L2 may substantially differ in terms of both quantity and quality of L2 input (Gass & Selinker 2008). Due to which, and as is empirically corroborated (Alsakran 2011; Deogratias 2011; Eftekhari & Rahimi 2011), the learning environment can be regarded as an important factor in the process of language acquisition. This also means that different learning settings, including micro-contexts such as university study programs within the EFL learning context, may have different effects on L2 collocation acquisition. It is thus hypothesized that the participants will show somewhat different collocation performance depending from which EFL learning environment they stem.

H6: Learners from two different 'EFL setting' groups will demonstrate different collocation performance in that one group will achieve better scores than the other.

1.4 Definition of terms

This section presents a list containing a number of key terms defined as they are used in the present study.

ESL: the term refers to the teaching and learning of English as a second language, implying that learners have daily access to an English-speaking environment outside the classroom context.

EFL: the term describes the teaching and learning of English as a foreign language where it is taught in a non-English-speaking region, i.e. in the students' home country, with no

exposure to the target language environment. In this case, the classroom constitutes the main source of exposure to English, also including personal out-of-classroom sources.

EFL classroom: the term stands for instructed learning that takes place within the confines of the classroom in a foreign language environment.

EFL environment: the term refers to the context of learning English as a foreign language in the learners' home country that occurs both in and outside the classroom. That is, in the EFL environment learners can be also exposed to English through other sources than the classroom, such as regular home reading, internet surfing, listening to music or watching TV in English.

Exposure: the term concerns all forms of contact with EFL/ESL, both inside or outside of instruction contexts.

Classroom/s: the term strictly refers to an instructional setting that is a formal context of learning the second/foreign language that occurs in schools and educational institutions.

Congruency: the term stands for the translational correspondence between L1-L2 collocations.

Congruent collocations: the term refers to combinations that have direct translation equivalents in the learners' L1.

Non-congruent collocations: this term concerns combinations that have no direct translation equivalents in the learners' L1.

Delexical/ized uses: the term refers to meaning senses of an item that go beyond its lexical one, as with the delexical verb *make* in the combination *make a complaint*.

Quantity of input (as a variable): the term refers to the length (years) of exposure to L2 input.

Quality of input (as a variable): the term broadly stands for the richness, authenticity, as well as the intensity (extra-curricular engagement) and type (classroom methodology, underlying ideology) of exposure to L2 input.

1.5 Thesis outline

The present study is organized into altogether five sections. The purpose of the next section (Section 2) is to provide a theoretical background to the empirical analysis. It consists of four major parts. In the first part, the term 'collocation' is introduced, including a brief account of different approaches to studying the phenomenon and a definition of the term as it is employed in the present investigation. This is followed by a description of the construct of the productive collocation knowledge to be measured as well as the role of collocations in language learning. The next two parts of Section 2 are dedicated to a review of relevant studies on L2 collocation knowledge and use and a discussion of L2 collocation acquisition,

which comprises an account of the general theoretical framework of collocation learning, factors influencing acquisition and a description of the EFL learning environment. Section 3 outlines the methodology employed in the thesis. In particular, it describes the setting, in which the study was carried out, the participants, the test instrument including analysis procedures, and ends with a summary of the research hypotheses. In Section 4, the results of the investigation addressing six research questions are presented and discussed. Finally, Section 5 draws conclusions of the thesis, starting with the significance of the study and an overview of the main findings and conclusions, followed by pedagogical implications. At the end, the limitations of the study are addressed along with some suggestions for further research.

2. Theoretical background and previous research

2.1 The phenomenon of collocation

2.1.1 Theoretical frameworks of collocation research

Collocations have been extensively studied and described under different terms, such as prefabricated units, prefabs, phraseological units, (lexical) chunks, formulaic sequences or multi-word units (Nesselhauf 2005: 1-4; Wray 2002: 9) that all broadly refer to the same phenomenon, namely to any meaningful association of words that occurs more frequently than expected by chance. ‘Lexical chunk’ is commonly used as an umbrella term for such word groupings. Despite multiple attempts at pinning down the notion of collocation, a theoretically more precise and commonly agreed definition of the term is still in progress. The existing terminological inconsistency is largely due to the highly complex and evasive nature of the phenomenon, which makes research on the topic often very difficult. Furthermore, the diverse conceptualizations and interpretations of collocation are significantly influenced by research agendas and, particularly, the underlying theoretical framework. Thus, in order to gain a better understanding of what is involved in collocation and how to define it, it is necessary to first take a look at its historical background.

To provide an introduction, the study of collocations can be said to have been started in the 1950s and since then it has been carried out under two distinct, though in some points overlapping, approaches that can be referred to as the ‘frequency-based approach’ and the ‘phraseological approach (Barfield & Gyllstad 2009), also known as ‘structural approach’ (Gitsaki 1999). Proponents of the two traditions focused on different aspects of collocation, using various methods. The approaches conspicuously differ in their conceptualization of the notion, concurring only in viewing collocation as representing some kind of syntagmatic lexical relationship. For the frequency-based approach, strongly associated with the fields of Corpus Linguistics and Computational Linguistics, statistics and frequency are at the heart of the investigation of textual realizations of collocations. Within the phraseological tradition, advocated mainly by linguists working in the area of Lexicography and Language Pedagogy, collocations are treated as a distinct type of word combination that is characterized by various degrees of fixedness and transparency. This approach is governed by syntactic and semantic analysis and goes back to Russian phraseology.

In particular, the frequency-based tradition originated in the work of Firth (1957), due to which the proponents of this approach are often referred to as Firthians. Firth’s attempt at the

description of a lexical meaning at the level of collocation was pioneering at the time, because it viewed meaning relations between words along the syntagmatic axis as opposed to the orthodox perspective of the paradigmatic relations (e.g. synonymy, antonymy). Firth's work was further developed by researchers interested in corpus and computation linguistics (Nesselhauf 2005). Advocates of this tradition generally define collocation as a sequence of co-occurring words at a certain distance from each other and a distinction is commonly made between frequent and infrequent combinations (Sinclair 1991; Halliday 1966). More specifically, collocations are viewed here as co-occurrences that are "more frequent than could be expected if words combined randomly in a language" (Nesselhauf 2005: 11–12). In addition, it was proposed that the probability of co-occurrence can be predicted by statistical measures (Halliday 1966). Thus, within the frequency-based approach, frequency of occurrence of items together is regarded as the key characteristic feature of collocation. The most prominent figures in this tradition are Firth (1957), Halliday (1966) and Sinclair (1987, 1991).

The conceptualization of collocation within the phraseological framework is strongly influenced by research conducted predominantly in Russia in the 1940s (Cowie 1998). To phraseologically-oriented researchers a collocation is a type of word combination that first of all exhibits degrees of fixedness and semantic transparency (Barfield & Gyllstad 2009). A further common ground for scholars of this tradition is the preoccupation with the description and development of classification systems for meaningful word sequences. For deciding whether a particular combination can be classified as a collocation, there is a tendency within this approach to adopt criteria based on specific lexico-syntactic and/or semantic characteristics exhibited by multi-words. Thus, unlike the frequency-based tradition, the phraseological school has been in principle less interested in frequency and statistical significance, but rather concerned with word combinations per se, i.e. their degree of semantic opacity and the substitutability of the lexical units in these combinations. As Howarth (1998a: 27) pointed out, the phraseological dimension is paramount and indispensable to the study of word combinations as insights from strictly mechanical statistical algorithms, though valuable, are often insufficient to account for such a complex phenomenon as collocation. Influential works within the phraseological framework have been carried out by Cowie (1998), Benson et al. (1997), Howarth (1996, 1998a, 1998b), Mel'čuk (1998), Nesselhauf (2005).

In addition to the two fairly crystallized frameworks, there can also be identified a third approach, to which Gyllstad (2007:15) refers as "the best of two worlds" because it tends to combine the virtues of the two lines of research discussed above. Thus, the third tradition can

be said to comprise approaches to collocation study that draw on both frequency-based and phraseological criteria. In particular, there is a growing tendency to utilize the frequency significance, the key criterion of the frequency-based approach, in addition to phraseological criteria, such as syntactic structure and/or semantic features of commutability and transparency (e.g. Altenberg 1993; Herbst 1996; Benson et al. 1997; Nation 2001; Nesselhauf 2003, 2005). Along these lines, a word combination constitutes a collocation if it is restricted in its substitutability, transparent in meaning and also occurs frequently. A representative of this approach is, for example, Nesselhauf (2003, 2005) who, primarily a phraseologist, employs frequency as an additional criterion in her analysis of the L2 use of verb-noun combinations. In particular, drawing on Howarth (1998a), she classifies multi-words into three categories (free-combinations, restricted collocations, and idioms) using the restrictedness criterion to discriminate between the groups and, on the other hand, the frequency of items in the British National Corpus (BNC) to establish acceptability of collocations produced by learners. Similarly, Stubbs (1995), basically viewing collocation as an association of habitually occurring words (either lemmas or word derivations), applies grammatical pattern as a complementary tool for analyzing collocations. Finally, the third approach involves accounts that theoretically do not easily fit into either of the approach schemas described. For instance, there are researchers who, while not explicitly using the term collocation, obviously deal with structures having all the hallmarks of collocations (Gyllstad 2007).

2.1.2 Definition of collocation in the study

As is clear from the above review, collocation is a complex phenomenon and attempts at capturing its definition and studying are multiple and diverse. Although two major paradigms can be said to have been developed (a frequency-based and a phraseological approach), there are points of considerable overlap between the two, with some scholars employing criteria from both traditions. For the sake of clarity, it is thus important to define collocation and describe the term as it is conceptualized and used in the present study.

First of all, it is necessary to distinguish between a collocation (countable noun) and the term collocation (uncountable noun). The former describes a type of combination of two or more words that occur at some distance from each other in written or spoken discourse. The latter term is used more generally to refer to the phenomenon of customary associations between lexical items. The term collocation is also used here to refer to an abstract linguistic unit having realizations in text, based on which the elements making up a collocation are

treated as lexemes. For example, the combinations *make a decision*, *makes a decision* and *made a decision* are assumed to represent instances of the same collocation. Moreover, following Nation (2001), collocations are addressed here from the word-property approach perspective. This means that collocations are conceptualized as a subcomponent of word knowledge interrelating with other properties of lexical competence (e.g. vocabulary, grammar, semantics, frequency or style) rather than representing an independent construct.

A theoretical approach to collocation adopted in the current study largely falls into the pattern of the third framework as both statistically-oriented and phraseological criteria are utilized to conceptualize and define the phenomenon. Following this, the term ‘collocation’ will be used here to refer to any conventionalized, habitually recurring lexical combinations exhibiting a certain grammatical structure, and that are to a certain degree restricted and semantically transparent. With regard to conventionality, it means that collocations are normally recognized by native speakers as appropriate and customary combinations. In other words, a collocation constitutes a meaningful, recurrent word group, which is attested in the present study by means of the BNC corpus as well as relevant dictionary sources. In terms of restrictedness, collocations are distinguished from free combinations whose constituents are easily interchangeable with other lexical items according to grammar rules. Semantic transparency refers to the feature of a collocation to express a meaning that is readily understood from the individual words that form them, which, on the other hand, distinguishes collocations from idioms whose meaning is often opaque. The restrictedness feature can be illustrated by the following examples: *money* customarily co-occurs with *big* but not with *large*, *conference* collocates with *hold* or *have* but not with *deliver*, and *speech* combines with *deliver* but hardly with *hold*. An example of semantic transparency is the following: *make in make a complaint* is not used with its original, lexical, meaning (e.g. *make a box*), but the phrase is clearer than *make a killing*, an idiom that means *make a large amount of money*. Collocations are mostly transparent in their meaning if learners know the individual words that compose them (e.g. *break a promise*, *take a risk*, *make a decision*).

In terms of the syntactic properties of collocations, following the phraseological tradition, in particular Benson et al. (1997), the present study also distinguishes between lexical (e.g. verb-noun, adjective-noun, adverb-adjective) and grammatical (e.g. noun-preposition, adjective-preposition) collocations, focusing on the former type, namely verb-noun. In addition, based on whether English collocations are direct L1 equivalents, i.e. readily (word-for-word) translated from their counterparts in learners’ L1, or rendered in a different pattern, involving at least one component that is different from the L2, the criterion of L1-L2 congruency is employed here to distinguish between two groups of combinations: congruent

and non-congruent. This classification is deemed to be more appropriate for the current cross-linguistic investigation than classification based on the criteria of semantic transparency or restrictedness, adopted by other L2 collocation researchers. For example, Nesselhauf (2003: 236), investigating non-native use of L2 verb-noun combinations (that she categorized into three types according to their restrictedness), made an observation that most restricted collocations - also most problematic for the learner- were non-congruent. In this case, it is thus difficult to say with certainty if it is the factor of congruency or restrictedness that caused difficulty in producing correct L2 collocations, because some restricted L2 collocations may have direct equivalents in learners' L1 and therefore be easier to produce, whereas combinations considered to be free combinable in L2 may be nevertheless difficult to translate into L1.

One possible explanation of this phenomenon is that collocations can be highly culture-specific, so that the same meaning is expressed by a pattern that may significantly differ across languages. This implies that phraseological criteria used for classifying English collocations cannot always be effectively applied to L1-L2 studies, as what is true of lexical patterning in one language cannot be the case in the other. To illustrate the point consider, for example, the situation where the restrictedness criterion proposed by Howarth (1998a) and Cowie (1998), based on which the word pair *make tea* is considered an unrestricted (free) combination, may not be applicable in relation to the Russian language, where there is only one verb equivalent (делать 'delat') for both English items *make* and *do*, often causing confusion when translated into Russian. That is, for the Russian learner, the English 'free' combination *make tea* most probably constitutes a restricted one as it is often erroneously translated into *do tea*, and therefore needs to be learned to use correctly.

Thus, it seems more reasonable to classify collocations based on their L1-L2 similarity, which is also acknowledged to be a crucial factor in collocation learnability (Nation 2001: 23-4). In fact, this approach has been already employed by some L2 researchers, however, not always explicitly described as such. For example, Gitsaki (1999) focused on the L2 use of collocations that she defined as restricted, in Howarth's (1998a) definition, but in addition to this it was also indicated that they are restricted in the sense of their non-congruency with the students' L1 (Greek). Consequently, given the inherent differences between languages in expressing meaning, it is more meaningful for an L2 collocation acquisition study to classify target collocations on the basis of their congruency with the learners' L1 equivalents rather than rely on properties initially proposed to describe collocations within the context of English as a native language, which, however, can be necessary when investigating learner population involving a mixture of L1s.

Thus, according to the definition used in this study, *buy a car* and *take a pen* are considered to be free combinations, *buy time* and *take steps* to be collocations, and *buy the farm* and *take a rain check* are regarded to be idioms. For the purposes of this analysis, collocations are also divided into congruent and non-congruent combinations depending on their similarity with the learners' L1 counterparts.

2.1.3 Collocation knowledge: defining the construct

In order to make the construct to be measured in the present study, i.e. productive collocation competence, clear and understandable, it first needs to be defined. This section is thus aimed to describe what is understood here under the productive level of collocation knowledge. The account largely draws on the vocabulary acquisition paradigm due to a lack of the established framework in the field of L2 collocation acquisition.

It is reasonable to start a discussion by defining the parts that make up collocations, i.e. words. It is not an easy task, however, because a word is a complex phenomenon that is part of a sophisticated, multi-layered system called language (Nation 2001). At a basic level, a lexical item can be defined as a sequence of letters separated by blanks as can be observed in written discourse (e.g. Milton 2009: 8). In this sense, the term 'word' basically stands for an orthographic word. This type of definition is useful if one is interested in how many words are, for example, in a text, or a student's essay. When words are counted in this way, they are often referred to as tokens. A common term for all the words in a language is vocabulary.

Despite this fairly straightforward explanation, in reality, however, the notion of vocabulary is much more complex than to be viewed in terms of the orthographic structure. The main problem with this definition is the fact that there is often no one-to-one correspondence between a form and a meaning, which means that a group of tokens can be used to express one meaning as, for example, in the case of multi-word units or collocations (e.g. *get a chance*, *make a complaint*). Moreover, in addition to the familiarity with the form and meaning of a word there is also a wide range of other aspects and nuances of knowledge necessary to use a word properly and in a variety of language contexts.

It is customary to describe vocabulary knowledge to include the following key components: meaning, written form, spoken form, grammatical behavior, register, frequency, associations, and collocations (Nation 1991). Thus, it can be seen that collocations constitute one of the properties necessary for understanding/using a word. From this description, however, it is not clear the exact status of this aspect of knowledge as well as how it interacts with other components, which is why a reference will be made to a more detailed descriptive

framework of word knowledge developed by Nation (2001: 27) (see Table 1 below). This description is considered to be one of the most elaborate ones so far (Gyllstad 2007: 34), which is why it will be used here to support the description of productive collocation knowledge for the present analysis. As shown in Table 1, three major categories are distinguished in this model: knowledge of form, knowledge of meaning, and knowledge of use.

Table 1. **Description of the word knowledge from Nation (2001: 27)**

Form	spoken	R	Can the learner recognize the spoken form of the word?
		P	Can the learner pronounce the word correctly?
	written	R	Can the learner recognize the written form of the word?
		P	Can the learner spell and write the word?
	word parts	R	Can the learner recognize parts in the word?
		P	Can the learner produce appropriate inflected and derived forms of the word?
Meaning	form and meaning	R	Can the learner recall the appropriate meaning of this word form?
		P	Can the learner produce the appropriate word form to express this meaning?
	concepts referents	R	Can the learner understand a range of uses of the word and its central concept?
		P	Can the learner use the word to refer to a range of items?
	associations	R	Can the learner recall this word when presented with related ideas?
		P	Can the learner produce common associations for this word?
Use	grammatical functions	R	Can the learner recognize correct use of the word in context?
		P	Can the learner use this word in the correct grammatical patterns?
	collocations	R	Can the learner recognize appropriate collocations?
		P	Can the learner produce the word with appropriate collocations?
	constraints on use (register, frequency)	R	Can the learner tell if the word is common, formal, infrequent etc.?
		P	Can the learner use the word at appropriate times?

R = receptive knowledge, P = productive knowledge

Each category comprises a set of knowledge types, including their receptive (marked as letters R) and productive (marked as letters P) aspects. Knowledge of form includes the spoken and written form as well as word parts. Knowledge of meaning involves form and meaning, concepts and referents as well as associations. Finally, knowledge of use consists of grammatical functions, collocations and constraints on use. Additionally, with the purpose of guiding of what is expected from a testee in relation to a particular word knowledge component, there are also questions in the fourth column pertaining to each aspect. The

proposed word knowledge framework has now a wide application in the context of ESL/EFL language learning and research, as it helps identify a particular knowledge component for investigation and testing purposes, as well as to understand how all the levels fit together in a complex system. As can be seen in the table, collocations pertain to the word use domain, and therefore can be regarded as rather sophisticated or deep knowledge of a lexical item as it implies fluent mastery of both syntactic (form) and semantic (meaning) lexical components, which is normally achieved at high levels of language competence.

Moreover, each word knowledge aspect, as is noted above, involves its receptive and productive competence. This distinction reflects the widely acknowledged view that vocabulary acquisition is not an all-or-nothing matter, but inevitably incremental. That is, a person learns a word in a gradual manner over a period of time achieving various degrees of mastery of various knowledge aspects. According to most models of acquisition, knowledge of words normally progresses from the receptive to the productive level. It is therefore a common practice in vocabulary research to distinguish between receptive or passive, and productive or active word knowledge (Gyllstad 2007: 35; Nation 2001: 24-5). Receptive knowledge typically refers to the ability to recognize and understand a word when listening or reading, while the capability of recalling and independently producing a word when writing or speaking is considered productive knowledge (Schmitt 2010: 87). It is generally believed that productive competence takes longer to achieve than receptive abilities as it is probably more difficult (Nation 2001: 28). In fact, there is empirical evidence that L2 learners are generally better at receptive than productive vocabulary mastery and that there is usually a significant gap between the two types of knowledge in favor of the former (Laufer 1998; Laufer & Girsai 2008; Schmitt 2010: 80).

Given the syntactic properties of collocations (i.e. composed of lexical items), there is no reason to believe that knowledge of a collocation, in particular with regard to the receptive-productive dimension, develops in a different way. It is possible to imagine a situation where an L2 user knows the meaning of the combination *make a complaint* and therefore can recall it when encountered in the input (receptive competence), but might not yet be able to produce it readily in writing or speaking. In other words, there can be levels of mastering collocations along the receptive-productive dimension as well. This logic, in particular, provides support for an approach to define collocation mastery drawing on the word knowledge framework, based on which, the following working definitions can be proposed. The receptive competence concerns the ability to recognize that a certain sequence of words confronted with in reading or listening constitutes an appropriate combination, including understanding of its meaning to some degree. For the productive collocation

competence, on the other hand, learners are expected to be able to recall and produce a conventional collocation in speech or writing. The relevant question here is: “Can the learner produce the word with appropriate collocations?”.

Moreover, there is empirical evidence showing degrees of knowing a word/collocation at a productive level as well. In particular, Farghal & Obiedat (1995) found differences in L2 learners’ performance on a translation task (free production) and a cloze test (prompted production) in favor of the latter one, which was probably easier. Similarly, Laufer (1998), who investigated the development of EFL vocabulary knowledge over a year of school instruction across three dimensions (passive, controlled active and free active), found significant gains only in passive vocabulary, followed by moderate progress in controlled active knowledge of vocabulary, while no tangible increase in free active vocabulary. Thus, drawing on this evidence, the present study also distinguishes between the controlled active and the free active aspects of the productive mastery of collocations.

To define the terms, Laufer’s (1998) description of the productive knowledge of words was adapted to the domain of collocations and the following definitions developed. Controlled-active knowledge of collocation (hereafter ‘cued production’) implies cued collocation recall, that is, the production of an item prompted by its collocate pair (including onset letters) as in a cloze test. Free active knowledge (hereafter ‘free production’) is defined here as referring to effortless and independent retrieval of a collocation, as expected in an L1-L2 translation task. It is hoped that such a two-fold distinction will allow a fine-grained analysis of the learners’ productive collocational knowledge using the corresponding test measures: a cloze test and translation. This approach, in particular, makes the study original to some extent as until now there have been no detailed discussions about learners’ nuanced productive collocational knowledge in current EFL collocation research.

2.2 The importance of collocations in language learning

One of the most significant and far-reaching changes in the field of Second Language Acquisition (SLA) in the late twentieth century was the shift from the emphasis on grammar (and phonology) to vocabulary, previously viewed as a secondary element or a mere by-product in language learning and teaching (Coady & Huckin 1997; McCarthy 1990; Nation 2001; Nattinger & DeCarrico 1992). Vocabulary is now widely recognized to play an important role in mastering a second language and as probably the essential language component for learners (Coady & Huckin 1997; Gass & Selinker 2008; Lewis 1993; Nation

1991; Schmitt 2000). Moreover, the increased interest in vocabulary along with the introduction of language corpora and corpus techniques gave rise to research that has begun to go beyond single words and focused on the analysis of multi-word units (e.g. Altenberg 1998; Nattinger & DeCarrico 1992; Sinclair 1991; Stubbs 1995; 2004).

A remarkable discovery from corpus-based studies has been a surprising amount of lexical patterning in natural language (Nation 2001; Schmitt 2000; Sinclair 1991). Such patterning indicates that vocabulary largely operates in chunks and clusters rather than individual lexical units. In other words, natural language was found to be largely formulaic, i.e. structured according to the idiom principle that hinges on collocations and other kinds of multi-words for expressing meaning and structuring discourse (Altenberg 1998; Biber et al. 1999; Sinclair 1991; Stubbs 2004). The main reason why phrasal expressions are so pervasive in a language is most likely because of a wide range of useful purposes they achieve in communication, e.g. functional use (e.g. speech acts as apologizing or requests), social interaction (e.g. formulas on agreeing with your interlocutor, providing backchannels), discourse organization, and precise information transfer (e.g. phraseology in various professional fields) (Schmitt 2010: 10-11). In this respect, knowledge of formulaic units can be said to offer processing advantages as it considerably reduces the cognitive load of analyzing and producing language from separate pieces rather than in ready-made chunks, thus leading to better language reception and production.

Growing evidence of the idiom principle in language use underscores the significance of syntagmatic or collocational links between words. Of all types of formulaic language, collocations have thus received especially much research attention, yielding numerous fruitful insights into the nature and behavior of this phenomenon (Barfield & Gyllstad 2009; Mel'cuk 1998; Nesselhauf 2005; Stubbs 1995). Collocation has been argued to constitute a key organizing element of the lexicon in particular, and an important principle underlying language structure and much of its patterning and relationships in general (Schmitt 2000). Currently, there is also strong evidence in support of the importance of collocations in language learning and use coming from an array of fields, such as cognitive linguistics (e.g. Langacker 1987), psycholinguistics (e.g. Wray 2002), corpus linguistics (e.g. Sinclair 1991), and research on first language acquisition (e.g. Tomasello 2003) and on oral fluency skills (e.g. Pawley & Syder 1983).

In particular, the contribution of corpus research has been in producing rich evidence of the omnipresence of collocations in natural language (McCarthy 1990; Carter 1998; Sinclair 1991; Siyanova & Schmitt 2008). By providing an enormous amount of natural language data, corpus has also played an important part in understanding the workings of the syntagmatic

dimension of a language as well as in describing and classifying collocations. Moreover, in addition to representing a strictly statistical phenomenon observable through a language corpus, there is currently a strong case for hypothesizing collocations to have also psycholinguistic validity, namely that collocations are, first of all, textual manifestations of the mental organization of words in meaningful strings processed whole (Hoey 2005; Schmitt 2010:141-2). In fact, there is empirical evidence suggesting that collocational associations represent strong links between words in the mind that seem to last until old age (Aitchison 2003). Thus, it seems that the human mind, rather than having two separate ‘files’ for storing grammar rules and individual words to be later combined to produce language, represents a complex network in which every lexical item is recorded with its collocate/s, so that one word automatically evokes a related one in a given context as, for example, when one hears the beginning of *home sweet _____* the word *home* will immediately come to mind.

Along these lines, collocations can be viewed as a useful mechanism that enables us to function in a language automatically, that is to process and produce language without needless cognitive effort. This kind of mechanism is present in almost every aspect of human behavior, such as walking, swimming, reading or riding a bicycle. For example, after having learned and extensively practiced to ride a bicycle, we finally arrive at a point where we no longer consciously process every move we make and how to make it, which in turn makes us ride faster and more skillfully. This example can serve as a good metaphor for explaining the function of collocations in a language, that is, to provide us with a (linguistic) skill enabling automatic (autopilot) operation of a routinized activity of speaking/writing. The practical application of this in relation to the learner is that mastery of a large number of common (as it is impossible to learn all) L2 collocations can help achieve considerable fluency and native-likeness in L2.

Cognitive benefits of using stored word combinations for language processing are currently widely recognized in the literature (Farghal & Obiedat 1995; Nation 2001: 335; Nattinger & DeCarrico 1992; Nesselhauf 2003; Wray 2002). Researchers believe that by making use of ready-made strings of words the language user can understand and structure a message without making a time-costly recourse to the underlying rules and thus having to recreate language from scratch. Processing advantages of knowing collocations leading to improved comprehension/production are also evidenced in a recent seminal study of Lin & Adolphs (2011) that is based on a spoken learner corpus and thus is a valuable contribution to previous L2 collocation research, mostly drawing on written learner data.

The significance of collocations as an essential language skill is further supported by the findings from language acquisition research that claims collocational knowledge to be

tantamount to language knowledge (Nation 2001; Ellis 2001). In this view, collocations are crucial because they are believed to form the basis of language knowledge and use in general. In particular, ready-made chunks are believed to represent stable linguistic patterns that are used as a raw building material for both understanding and constructing language, namely without the step of referring to underlying rules. Most importantly, knowledge of collocations is argued to drive not only the process of L1 acquisition (Ellis 2001; Peters 1981; Tomasello 2003; Wray 2000), but also second/foreign language learning (Bardovi-Harlig 2002; Nattinger & DeCarrico 1992).

Conversely, insufficient knowledge of collocations may considerably impede L2 competence both receptively and productively, leading to overall disfluency and non-idiomaticity. As Hill (1999: 5) pointed out, “students with good ideas often lose marks because they don’t know the four of five most important collocations of a key word that is central to what they are writing about”. As a consequence, they produce longer sentences filled with monotonous and repetitive utterances, failing to convey precisely what they intend to say. Moreover, poor collocation competence can serve as a strong sign of non-nativeness among language users as the use of phraseology may signal group membership of a particular discourse community (Nesselhauf 2005; Pawley & Syder 1983; Wray 2002). This is why collocation mastery can also facilitate the integration into the target linguistic community (Pawley & Syder 1983).

All in all, the above discussion serves to illustrate that a good command of collocations is a crucial component of L2 competence. This, in turn, makes a strong case for dedicating a great deal of teaching efforts and classroom time to collocations if learners are to attain native-like fluency and idiomaticity in both the spoken and written modes of an L2.

2.3 L2 knowledge and use of collocations

Although it is widely accepted that collocations play a major role in language learning and teaching, research into non-native speaker (NNS) knowledge and use of L2 collocations, in particular in the context of EFL, still remains relatively scarce. In this sub-section, a general overview of studies on L2 collocation knowledge and use will be provided to outline what is known to date about the subject, focusing on works that are particularly relevant to the present thesis. The survey is limited to studies of learners of English in which target multi-word units are considered to be collocations in the present sense of the word. Overall, research on L2 collocations is characterized by a wide diversity in terms of theoretical

assumptions, methods, measures and types of target items, as well as subjects' L2 proficiency and L1 backgrounds. With regard to methods, two major approaches can be clearly identified: analysis of collocation production in a learner corpus and the adoption of various collocation elicitation techniques. For ease of reference, the overview is structured around these two main research designs.

2.3.1 Corpus-based research

Owing to the recent advances in computer technology, a growing number of studies employ corpus methodology to investigate different aspects of L2 collocation knowledge and use, e.g. Altenberg & Granger (2001), Chi et al. (1994), Gitsaki (1999), Granger (1998), Fatemeh & Arshad (2011), Howarth (1996), Kaszubski (2000), Laufer & Waldman (2011), Nesselhauf (2005), Siyanova & Schmitt (2008) and Zhang (1993). From a theoretical perspective, corpus studies can be characterized by adopting a predominantly phraseology-oriented view of collocation as they tend to classify word combinations based on their various semantic and/or structural properties. The adoption of corpus analysis techniques has the benefit of collecting a large variety of data for analyzing L2 use of different types of collocations and on a wide range of dimensions: accuracy/appropriacy of use, overuse or underuse of specific collocations or their components, as well as for comparing learner performance with native-speaker production. There is a strong tendency among researchers, however, to concentrate on overuse or underuse by comparing the amount (tokens) of particular collocation types produced by non-natives against the native-speaker norms. The exceptions are those studies that are interested in the actual L2 use of collocations (e.g. Chi et al. 1994; Nesselhauf 2005) or their development (Gitsaki 1999).

One of the earliest studies in the field is that of Chi et al. (1994), who analyzed a one-million word corpus of essays written by Chinese L1 first-year students in terms of possible deviations in the use of L2 collocations containing common delexical verbs (*do*, *get*, *have*, *make*, and *take*). The main result of the study was that learners often experienced difficulty in knowing which verb belonged with which noun, erroneously using the verbs as if they were interchangeable (e.g. combining *have* with *decision*, *progress*). Chi et al. thus concluded that combinations involving common delexical verbs might be especially problematic for NNSs. The same set of delexical verbs (including *be*) was investigated by Kaszubski (2000). Common recurrent combinations involving these verbs were extracted from a corpus of English native speakers (NSs) and compared to their non-native use in the corpora of different EFL learner groups (intermediate Polish and Spanish L1 learners, advanced Polish and French

L1 learners). The learners were found to use overall fewer collocations than NSs but to consistently overuse a small range of favorite ones, especially those that either had a high frequency in English or similarity to L1. The findings of these two studies are largely consistent with a more recent investigation of Altenberg & Granger (2001), who examined the EFL learner use of lexical combinations involving high-frequency verbs, focusing on a major representative of the group *make*. The study compared the collocation patterning of *make* across three corpora: two corpora of learners with Swedish and French as L1 and a NS corpus. The results indicate that EFL learners, even at an advanced competence level, experience considerable difficulty with high frequency verbs like *make*. Particularly problematic proved to be delexicalized uses of the common verbs like *make*, *get*, *take* (e.g. *make a decision*, *make a complaint*) as they enter a large number of formulaic strings and collocations expressing various meanings.

A further investigation of L2 use of collocations as compared to NSs was conducted by Howarth (1996) in one of the most rigorous studies of collocations in learner language to date. Howarth manually analyzed verb-noun collocations in essays written by 10 advanced English students with 8 different L1s in a corpus of English academic writing (about 23,000 words). He distinguished between three major categories of multi-word units: *free combinations* (e.g. *buy a house*, with both parts freely combinable), *restricted collocations* (e.g. *take a picture*, where the collocability of the verb is restricted), and *idioms* (e.g. *kick the bucket*, where substitution of both constituents is limited). When compared to the NS corpus, it was found that L2 learners used more free combinations, but about 50% less restricted collocations and idioms. A large proportion of atypical verb-noun pairs in learner data was due to the production of combinations by way of blending two collocates with similar semantics, as in the case of **perform a project* assumed by NNSs to be acceptable based on the similar combination *perform a task* (ibid.1996: 37). It was concluded that out of the three collocation types restricted collocations seem to pose a particular challenge to learners. Finally, the study found a very low correlation between the use (number) and acceptability of restricted collocations and learners' general level of L2 proficiency. However, the generalizability of this result can be subject to question given the comparatively small number of the participants in the study (10).

Nesselhauf (2005) conducts another large-scale investigation of NNS use of verb-noun pairs categorized into free combinations, collocations and idioms, drawing on Howarth's classification. Contrary to Howarth, who compared learner language with NS data, Nesselhauf focuses on the actual L2 production with the aim of identifying the nature of the problems learners experience in dealing with collocations. The analysis was based on a corpus

of essays (around 155,000 words) written by advanced German L1 learners of English. A number of important results were produced. As for the overall learner production, it was found that a significant proportion (one third) of L2 combinations was non-native like or deviant, with the verb element most frequently being misused. Nesselhauf thus concludes that verb-noun collocations constitute a serious area of difficulty for learners, even at advanced levels. With regard to factors found to correlate with difficulty, it emerged that, first of all, L1 congruence played a major part, followed by the degree of restriction of collocations. Two further important findings were that 1) the length of classroom instruction did not show to significantly contribute to collocation competence, whereas the length of L2 natural exposure (stays in an English-speaking environment) was found to have a tangibly positive effect. The latter finding, in particular, emphasizes the role of learning environment in L2 collocation acquisition.

One recent comparative study on the L2 use of verb-noun collocations was conducted by Laufer & Waldman (2011). For the analysis, a large L2 learner corpus (about 300,000 words) was compiled consisting of essays written by Hebrew and Arabic speaking learners of English (college and university students) divided into three proficiency levels (identified as basic, intermediate and advanced). As a comparison baseline served a corpus of young adult native speakers of English. The investigation differs from other studies in that it performed two types of comparison: 1) each learner group was first compared on the frequency of collocation use to their English counterparts, and then 2) L2 collocation performance as regards frequency and correctness was compared between and within the learner groups. The first analysis revealed that learners at all proficiency levels significantly under-produced collocations in comparison to NSs of a similar age. This result is in line with the findings of some earlier studies (e.g. Howarth 1996; Kaszubski 2000), except for Siyanova & Schmitt (2008) who reported L2 learners to be on the whole as capable of producing appropriate collocations and in the same amount as native speakers. Furthermore, only advanced learners showed a significant increase in the percentage of collocations (tokens) when compared to the basic and intermediate learner sub-corpus. As regards correctness, it was found that learners at all levels of proficiency equally produced a fairly high proportion of non-standard collocations. A within-group comparison, however, surprisingly revealed that intermediate and advanced learners produced a considerably higher number of atypical collocations than the basic learners. This finding indicates that not only does the number of errors not decrease at a higher proficiency level, but as proficiency increases, the proportion of errors increases in parallel. Lastly, about half of all erroneous collocations produced were L1-based and did not seem to reduce with time. Overall, the results of the study suggest that collocation acquisition

is a particularly hard and uneven process that may require a long time and a special teaching/learning effort.

Granger (1998) also compares native and non-native collocation competence focusing, however, on adverb-adjective combinations consisting of *-ly* amplifiers (e.g. *completely*, *totally*) based on an L2 corpus (250,000 words) of advanced French L1 learners and NS essays. The overall findings correspond to the results for verb-noun collocations reported in a number of other studies (Howarth 1996; Kaszubski 2000; Laufer & Waldman 2011; Nesselhauf 2005) showing a significant underuse of native-like combinations by learners and adherence to non-standard combinations. Most preferred word pairs were often direct L1 translation equivalents and/or those that were less restricted on collocability. An interesting observation was that NNSs tended to overuse a small set of favorite intensifiers (e.g. *completely* and *totally*) used as “safe bets” (ibid.1998:148). Granger came to a conclusion that learners’ production of combinations considered unacceptable by natives is most likely due to a lack of a sense of salience for collocations in the target language.

In contrast to the majority of corpus-based studies, a number of investigations pursued aims other than the L2 overuse/underuse of collocations, concentrating instead on such aspects as the development of L2 collocation knowledge, as well as on the relationship of collocation mastery to other language skills or general proficiency in the L2. Given their relevance to the present discussion, I will concentrate here on the works of Zhang (1993), Fatemeh & Arshad (2011) and Gitsaki (1999).

Zhang (1993) explored the relationship between L2 collocational competence and writing quality. In the analysis, a writing task and a 50-item blank-filling test, based on 21 target collocations, was employed. Collocation performance of 30 college freshmen students of English with mixed L1s, and 30 NSs was examined by performing between- and within-group comparisons. The main results indicate that collocational knowledge indeed contributes to more natural-like and fluent written communication. Similarly, Fatemeh & Arshad (2011), also using a combined methodology (a writing task and a variant of a cloze test) examined the relationship between learners’ collocation knowledge and productive skills in the L2, in particular writing. The study focused on the use of verb-noun combinations in a corpus of stories produced by 27 PhD Iranian students (native speakers of Persian) over a period of six weeks. The results showed a high correlation between the students’ collocational competence and the use of collocations in their writing.

Gitsaki (1999) undertook a large-scale investigation designed to explore developmental patterns of L2 learners’ collocational competence. The study is notable for providing new and valuable perspectives onto the area of L2 collocation knowledge and use. The informants of

the study were teenage Greek L1 learners divided into three English proficiency groups (post-beginners, intermediate, and post-intermediate). Three measurement tools were employed: a writing task, a translation task and a cloze task. The learners' essays were analyzed as regards the occurrence of 37 different types of both grammatical and lexical collocations. The main finding of the study was that accuracy in the use of collocations increased in parallel with the learners' improved proficiency in the L2. Of all three proficiency groups, the post-intermediate students were reported to be most accurate and also most diverse in the production of both lexical and grammatical collocations. Among all item types, verb-noun collocations were found to be most difficult to translate and also to infrequently occur in the learner corpus, which supports the findings of other studies reporting verb-noun collocations to present a serious challenge to the L2 learner (Howarth 1996; Laufer & Waldman 2011; Nesselhauf 2005). Taken together, the above three studies, contrary to Howarth (1996), suggest that the use of acceptable collocations and of a large number significantly correlates with general L2 proficiency.

2.3.2 Elicitation studies

In order to investigate ESL/EFL learners' collocation knowledge, numerous researchers designed special collocation tests. Comprehension of collocations has been mainly studied by means of multiple-choice questions or acceptability judgment tasks, while production has been elicited using either fill-in-the-blank tests (a cloze test) or translation tasks or both. Since comprehension is generally agreed to be less demanding for the language user (Nation 2001: 28; Schmitt 2010: 80; Laufer 1998), there is a growing tendency among elicitation studies to concentrate on the productive side of L2 knowledge of collocations.

Researchers that focus on receptive collocation competence are interested in the learners' ability to identify whether certain combinations are appropriate in L2. Granger (1998), for example, examined (French L1) EFL learners' receptive knowledge of English combinations in addition to the analysis of their written production, reviewed earlier. A comprehension test designed for the study consisted of 11 amplifier adverbs and a list of 15 adjectives to be combined. The results of the study showed that L2 learners marked significantly fewer acceptable combinations than their English NS counterparts (280 versus 384), thus demonstrating a "weak sense of salience" of L2 lexical relations (ibid. 1998:152).

A notable investigation using a receptive collocation measure was carried out by Siyanova & Schmitt (2008), who conducted a series of three studies to examine NNS collocation production as well as the fluency and underlying intuitions of this production. In

particular, Study 1 analyzed collocation production in an (Russian L1) EFL learner corpus. Studies 2 and 3 intended to explore the mental processing on which that production is based as well as the speed of that processing as compared to NSs. The participants in the two studies were a mixture of advanced ESL and EFL learners (both students and professionals) with 9 different L1s. The main finding was that, despite showing a comparatively good productive ability in the first study, the learners' underlying intuitions of collocations were proved to be much weaker than those of native speakers. With regard to processing, NSs were again found to be significantly faster than L2 learners. Furthermore, evidence was provided that extensive exposure to high quality L2 input (through long stays in an L2 environment) considerably improves learners' intuitions about L2 collocations.

Among more recent receptive studies worth mentioning here are those by Eftekhari & Rahimi (2011) and Yamashita & Jian (2010). Eftekhari & Rahimi (2011) examined L2 collocation mastery and its relation to overall language proficiency using an acceptability judgment test. The analysis focused on the L2 use of 64 non-congruent English collocations involving the common verbs *have*, *give*, *take* and *make* in their delexical uses. The participants were 45 Iranian EFL students from three proficiency levels: low, intermediate and advanced. The results of the test revealed the learners' overall poor receptive mastery of collocations. Moreover, it was observed that errors in collocations seemed to fossilize at an intermediate level, showing a high resistance to change through an increase in L2 proficiency. This finding thus corresponds to Howarth's (1996) study, also reporting weak L2 proficiency effects on collocation knowledge. Yamashita & Jian (2010) used an acceptability judgment task to investigate the L1 influence on the acquisition of verb-noun and adjective-noun collocations by Japanese L1 learners of ESL and EFL. Overall, the results of the study showed that both the amount of L2 input and L1-L2 congruency can have a positive effect on the collocation acquisition process. Specifically, ESL learners (with more extensive L2 input) were found to significantly outperform EFL users, both in terms of accuracy and speed. Furthermore, the study indicates that non-equivalent L1-L2 collocations are not readily acquired even via extended exposure to L2, however, once retained, they are amenable to the procession independent of the learners' L1.

In addition to the above studies, there were also researchers that investigated both receptive and productive L2 collocation knowledge pursuing varying aims, for instance Bonk (2001) and Alsakran (2011). Alsakran (2011) examined the knowledge of 48 grammatical and lexical collocations (verb-noun, verb-preposition and adjective-noun) by ESL speakers as compared to EFL learners, using a gap-filling test and an acceptability judgment test. The results of the study revealed a considerable difference between the learners' productive and

the receptive collocation knowledge that by far exceeded the former one. Furthermore, ESL learners were found to significantly outscore their EFL counterparts, thus indicating that a learning environment plays a significant role in collocation acquisition. Bonk (2001), on the other hand, employed a mixture of a receptive and a productive (fill-in-the-bank) test of (50) verb-related collocations of various types to examine the reliability of the test instrument and a relationship between learners' (with mixed L1s) collocation competence and general English proficiency. The results demonstrated a high legitimacy of the collocation test as well as a significantly strong correlation between general L2 proficiency and learners' collocation mastery, which is again in contradiction with Howarth's (1996) study that found no positive correlation.

As mentioned earlier, most elicitation studies on L2 collocation knowledge were interested in investigating non-native collocation production, using either a translation task or a cloze test or both. Investigations carried out along these lines that will be briefly reviewed here, given their current relevance, are those by Aghbar (1990), Al-Zahrani (1998), Bahns & Eldaw (1993); Biskup (1992), Farghal & Obiedat (1995), Gitsaki (1999) and Hasselgren (1994). To begin with, it should be noted that on the whole the results from research on productive collocation knowledge are largely consistent with the findings from receptive tests that revealed L2 learners' overall inadequate syntagmatic competence.

A methodologically notable study conducted by Biskup (1992) investigated the ability of Polish (34) and German (28) L1 advanced learners to produce English verb-noun and adjective-noun collocations, employing a translation task. The results of the analysis showed that the Polish students produced altogether more collocations and, at the same time, much more frequently provided no answer at all than the German students, who made more errors by trying alternative ways of rendering the meaning of collocations whose L2 equivalents they obviously did not know. Biskup attributed the results to different emphasis in the education systems in Poland and Germany. Specifically, the Polish education system puts great emphasis on accuracy in EFL pedagogy, so the Polish learners showed avoidance in providing a translation unless they were sure it was the correct one. The German education system, on the other hand, focuses on developing fluency and communicative skills, which thus explains the German learners' creative strategy in supplying a definition or paraphrasing instead of avoiding translating unfamiliar collocations. As regards the error types in learner data, it was observed that L1 had a strong adverse effect on collocation performance of all the students; however, different types of L1 transfer were preferred by learners in the two groups. Considered together, the results suggest that taking into account such factors as the emphasis in EFL/ESL instruction can yield new perspectives on the issue of vocabulary acquisition. In

particular, the methodological value of Biskup's study lies in highlighting the importance of employing different approaches when studying L2 collocations.

Valuable insights into L2 learner collocation use were also produced by Hasselgren (1994), who investigated nonnative production of adverb-adjective and verb-noun collocations based on longer translation samples. Advanced Norwegian EFL learners participating in the study were asked to translate texts (about 2,000 words each) which were then analyzed in terms of deviant uses as compared to the native speaker norms. Hasselgren found that the learners tended to overuse unrestricted intensifiers such as *very*, and core verbs like *get* or *give*. An overall important observation made in the study was that, when dealing with L2 collocations, non-natives seem to largely depend on familiar structures, termed by Hasselgren as lexical "teddy bears", i.e. items that are close L1 cognates, learnt early and/or used frequently, or simply favorite. A further major strategy was proved to be word-for-word translations from L1 equivalents, however, often resulting in erroneous/unacceptable L2 combinations.

In contrast to the two above mentioned studies, Aghbar (1990), Bahns & Eldaw (1993) and Al-Zahrani (1998) employed only a cloze test to explore L2 collocation production. Aghbar (1990) compared native and nonnative knowledge of 50 verb-noun collocations embedded in a sentence appropriate for formal written contexts. The performance was compared across the three learner groups: 27 university teachers of English, 44 American undergraduate students and 97 advanced university ESL students. The results showed that the faculty members outperformed the other two groups in collocation performance both as regards frequency and the number of correct responses. The American students demonstrated a good command of common collocations, but difficulty with combinations of a more formal style, whereas ESL learners showed the least developed collocation knowledge by providing the lowest number of correct responses. ESL students were reported to perform well only on the combinations where the most common verb *get* was involved, also showing a tendency to overuse it where more specific verbs such as *acquire*, *gain*, *obtain* were required. This finding is thus in line with previous studies that also reported L2 learners to rely on a limited number of favored collocations (e.g. Granger 1998; Hasselgren 1994; Kaszubski 2000).

Likewise, Al-Zahrani (1998) explored L2 productive collocation mastery, however, of (81) EFL Saudi students across four academic years (freshmen, sophomores, juniors, seniors). Additionally, the relationship between the informants' collocation competence and general L2 proficiency was analyzed. A total of 50 verb-noun combinations were tested in a cloze test, containing gaps to elicit the missing verb collocate. The results revealed substantial differences in collocation mastery between the students in the first two (freshmen and

sophomores) and the last two academic years (juniors and seniors), while almost no difference between the second-year and the third-year students, thus suggesting that collocation acquisition is a complex process that might take a long time to show substantial progress.

Bahns & Eldaw (1993) used a cloze test in addition to a general vocabulary knowledge test to investigate a relationship between the learners' productive knowledge of (15 verb-noun) collocations and general lexical proficiency. The participants were 58 German L1 university students of English, in years 1-3, which were equally divided into two groups to either provide L1-L2 translation of sentences (vocabulary test) or to complete a cloze test. The study found no significant differences between the groups in terms of the proportion of correct collocations. It was observed, however, that twice as many errors were produced in the translation of collocations as in the translation of individual lexical words, and that nearly 52% of non-standard collocations were produced in the cloze test. Based on these results, it was concluded that collocation competence is particularly hard to achieve and that it does not seem to develop alongside general vocabulary knowledge.

Finally, there are researchers that employed both a translation task and a cloze test, for example Farghal & Obiedat (1995) and Gitsaki (1999). Farghal & Obiedat (1995) intended to examine EFL learners' strategies in producing 22 common lexical collocations. The participants were 57 Arabic L1 university students of English divided into two groups to do either a fill-in-the-blank task (Group A) or a translation task (Group B). The results showed that Group A produced a much higher percentage of correct collocations than group B (18% to 5%). In terms of the strategies, the use of synonymy was found to be the most preferred one for both groups. The other two major strategies identified were translation from L1 equivalents, followed by paraphrasing. Interestingly, the results from the two types of test are suggestive of the impact a particular elicitation technique might exert on the strategy learners use in dealing with L2 collocations. Specifically, paraphrasing seems to be particularly encouraged in a translation task. Reliance on, or, more precisely, interference of L1 also tends to be stronger when L1 collocations are to be translated into L2. The main conclusion from the study is that L2 learners are on the whole deficient in collocation knowledge because "they are not being made aware of collocations as a fundamental genre of multi-word units" (ibid: 326).

Finally, Gitsaki's (1999) study, based on the same 275 learners reported in the above account of her essay-based investigation, employed a) a translation task consisting of 10 sentences with collocations of (6) various types to be translated from Greek to English, and b) a fill-in-the-blank test targeting eleven types of collocations presented in a one-sentence context. Care was taken to include restricted collocations in the sense of their non-congruence

with respective Greek equivalents. The main finding from the elicitation tests indicates that collocation competence develops as L2 learners' overall language proficiency increases. In particular, post-intermediate level students were found to demonstrate more accuracy in the production of collocations than the lower proficiency groups. Also, as was reported earlier, of all the collocation types tested, verb-noun pairs were found to be most difficult for the learners both in translation and the cloze test regardless of the L2 proficiency level. Gitsaki attributes non-native difficulty with collocations to their inherent features, namely arbitrariness and unpredictability.

2.3.3 Summary of the reviewed studies and key findings

For ease of reference, the key features as well as major findings of the reviewed studies are summarized in Appendix I. From the preceding account, a number of trends in L2 collocation research can be identified that are of importance to the present investigation. Firstly, non-native mastery of English collocations has been addressed along at least four dimensions: amount of use (contrastive analysis), accuracy/appropriacy of use (error analysis), relation to overall target language proficiency (inclusion of proficiency tests), and intuitions of L2 collocations (fluency/acceptability judgment task).

Secondly, two major study designs have been employed: corpus-based research, exploring learners' free collocation production, and elicitation studies that investigate learners' receptive and/or productive knowledge, concentrating on the latter knowledge type considered more problematic for non-natives. The majority of corpus-based investigations examined L2 collocation production focusing on underuse/overuse or deviant use as compared to the native speaker norms, whereas only few studies explored the actual non-native collocation knowledge or its development (Chi et al. 1994; Gitsaki 1999; Nesselhauf 2005). Most elicitation studies employed a cloze test and/or translation to tap into learners' productive knowledge of particular types of collocations, with only a small number of investigations using a receptive recognition task to test learners' collocation comprehension and/or intuitions (Eftekhari & Rahimi 2011; Siyanova & Schmitt 2008; Yamashita & Jian 2010).

Thirdly, as regards the L2 learner profile, in most studies, the participants were either only EFL speakers or ESL learners. In many cases, a learner sample also constituted a mixture of these two populations. The exceptions are the studies by Alsakran (2011) and Yamashita & Jian (2010), who actually investigated collocation competence by ESL users as compared to EFL learners, revealing considerable differences in performance between the two groups in

favor of the former one. Also, learners with mixed L1s often tended to be used to represent one single L2 learner population (e.g. Bonk 2001; Howarth 1996; Kaszubski 2000; Siyanova & Schmitt 2008; Zhang 1999). This, however, can be regarded as a serious methodological shortcoming as ample evidence indicates learners' L1 to be a significant factor in non-native collocation knowledge and use (e.g. Schmitt 2010: 25-26; Nesselhauf 2005; Yamashita & Jian 2010). In general, disregarding the individual L2 learner profile in researching non-native collocation knowledge, so that speakers from different L1 backgrounds as well as learning contexts are lumped together in a sample, might considerably decrease the generalizability of results.

What is more, research examining learners with Russian/Ukrainian as L1 are conspicuously absent in the field. Of all the reviewed studies only Siyanova & Schmitt (2008) have included Russian L1 learners of English, while other researchers investigated non-natives mostly with French, German, and Arabic as their mother tongue, followed by Polish, Greek, Swedish, Chinese or Spanish. However, although Russian L1 learners were part of the learner sample in Siyanova & Schmitt's (2008) study, the role of the L1 on the participants' performance was disregarded in the discussion of the results. This means that so far no studies have been carried out that focused on Russian L1 learners' collocation knowledge and use.

Lastly, among all combination types, lexical verb-noun collocations were most extensively studied, followed by the adjective-noun type. The reason for this might be, first of all, the importance of such collocations in conveying core meaning in a sentence, and also the fact that they pose a considerable challenge for L2 learners, particularly restricted items or those involving delexicalized verbs. Only a few studies, however, explicitly focused on the analysis of the L2 use of non-congruent collocations or those involving common verbs in delexical uses (Altenberg & Granger 2001; Chi et al. 1994; Gitsaki 1999; Kaszubski 2000; Yamashita & Jian 2010). Two other collocation types that were also often investigated are adverb-adjective and adjective-noun combinations.

As regards the main findings and conclusions, research consistently showed that collocations constitute a major stumbling block for L2 learners, regardless of their mother tongue, the amount of instruction they received in L2, general L2 proficiency or the type of language task they were subjected to. In particular, in comparison to native speakers, learner language was discovered to be less idiomatic due to a considerable underuse of collocations and, on the other hand, overuse of a limited range of favored ones, either because they were L1 congruent, learnt earlier, or highly frequent in the input (e.g. core verbs as *be*, *have*, *make*, *get* or amplifiers *very*, *completely*, *highly*). In terms of the receptive collocation competence, although Siyanova & Schmitt's (2008) study provided general counterevidence in support of

the learners' ability to produce a large number of acceptable combinations, underlying intuitions of collocation production, however, was found to be fairly underdeveloped.

Moreover, learner collocation performance was reported to be replete with erroneous/deviant combinations. A large proportion of collocation errors in most studies were L1-induced (interlingual) and persistent even at advanced levels (Biskup 1992; Farghal & Obiedat 1995; Laufer & Waldman 2011; Nesselhauf 2005). Among other non-L1-related (intralingual) errors attested were over-extensions of unrestricted verb uses as with *make* (e.g. **make an experiment*), or blends of semantically related collocates such as **pay effort* and *pay attention*. Most problematic for the learner proved to be collocations involving common verbs, especially in delexical uses (e.g. *make, take, get*), as well as non-congruent and restricted collocations (Altenberg & Granger 2001; Chi et al. 1994; Gitsaki 1991; Kaszubski 2000; Nesselhauf 2005; Howarth 1996; Yamashita & Jian 2010).

What remains inconclusive, however, is whether L2 collocation competence is related to general L2 proficiency. On the one hand, it was found that the knowledge of single words did not seem to develop in parallel with collocation knowledge (Bahns & Eldaw 1993), which largely corresponds with other studies revealing no positive (a weak) relationship between collocation mastery (correctness) and general competence in L2 (Eftekhari & Rahimi 2011; Howarth 1996; Laufer & Waldmann 2011). On the other hand, there are findings indicating a close relationship between the variables (Gitsaki 1999; Bonk 2001; Fatemeh & Arshad 2011; Zhang 1993). One explanation of such contradictory results can be probably the fact that the studies focused on different collocation types (in terms of their inherent difficulty), also using differing methods, which altogether might have significantly affected learner production.

For example, if one takes a closer look at the studies representing two opposing views, the following picture emerges. The studies that found a positive correlation explored a wide range of different collocation types, in particular Gitsaki (1999) used 37 different types (grammatical and lexical) of restricted combinations, Zhang (1993) 21 types, both restricted and unrestricted, and Bonk (2001) over 3 different types of combinations involving a verb. However, a careful examination of the other group of studies, claiming no positive relation, reveals, on the other hand, a narrower focus on the L2 use of high-learning burden collocations, namely verb-noun pairs, especially that are restricted/non-congruent ones (Bahns & Eldaw 1993; Laufer & Waldman 2011) and also those involving verbs in delexical uses (Eftekhari & Rahimi 2011; Howarth 1996). From this, it can be concluded that the success with which L2 proficiency contributes to collocation mastery is highly dependent on the type/difficulty of items, because some collocations, in this case the restricted/non-congruent verb-noun type, might be more difficult to learn than others.

From a critical perspective, it can be stated that L2 collocation studies, although having produced valuable insights into the subject, are not altogether without shortcomings. Despite the fact that most investigations focused on the intermediate to advanced learner, they vary widely as regards methods of enquiry, types and difficulty of target collocations as well as L2 learner profiles. Also, there is a great variation in size and/or scope of the investigations. In some elicitation studies, particularly early ones, 10-20 collocation pairs were tested, in some cases 50, while some production studies were based on a limited number of collocations or essays. This can be considered a flaw in a research design as with a small number of target items or L2 learners it is hard to draw well-grounded conclusions and generalizations (Barfield & Gyllstad 2009: 12). Great variation is also observed in item selection criteria. For example, for her analysis, Gitsaki (1999) selected target collocations from the participants' English textbooks used in their (Greek) schools. Bahns & Eldaw (1993) drew on English textbooks and dictionaries, while Fraghal & Obiedat (1995) chose a set of common combinations related to core topics such as food, colors and weather. Yet, there are several studies that do not discuss how their target items were selected at all (Biskup 1992), which can be viewed as a serious methodological omission.

In a nutshell, as the account above shows, L2 collocation research is characterized by a great diversity in terms of theoretical assumptions, research procedures, and learner groups, which makes it increasingly difficult to draw valid conclusions and generalize about the findings to build a single theoretical framework on L2 collocation knowledge and use. At the same time, such diversity can be a positive sign of the developmental stage of this rapidly increasing field that welcomes new investigations to corroborate the existing findings and help expand knowledge about L2 collocations.

2.4 L2 collocation acquisition

2.4.1 Placing into a broader framework

Despite a growing recognition of collocations as being crucial for fluent and idiomatic language use, the mechanisms of how they are acquired are, however, still not fully understood and explored (Schmitt 2010). As the review of previous studies also demonstrates, research in the field has primarily focused on what EFL/ESL learner know and are able to produce, in particular shedding some light on certain areas of difficulty learners experience in dealing with L2 phraseology. Thus, there is an obvious lack of an established framework within which non-native collocation acquisition can be interpreted and described. It has been

proposed, for example, that acquisition of formulaic language can be explained from a connectionist, or exemplar-based/usage-based view (Schmitt & Carter 2004; Tomasello 2003). According to this model, learning does not rely on an innate language faculty or on rule systems, but occurs via the extraction of regular patterns from a language environment. In this view, knowledge of a language is thus regarded as a complex network of patterns and structures that language users create as they encounter recurring linguistic forms. The emphasis is on usage, in that the more often elements are used together, the stronger they become interconnected. Frequent exposure to co-occurring words is believed to be crucial for creating and strengthening collocational associations between lexical items. From this paradigm perspective, learning thus takes place as language users extract patterns - items/words experienced together - to which they are regularly exposed in the input.

A similar, often referred to, model of collocation acquisition in a first language has been proposed by Ellis (2001). This model postulates collocation learning to be one level of chunking, i.e. storing in long-term memory of meaningful sequences of elementary (phonemes) to more complex constituents (such as words) that occur at all levels of language both in the written and spoken mode. Ellis explains that chunking, similar to exemplar-based learning, occurs according to the principle of associative pairing, following the power law of practice. That is, the same linguistic units experienced together in the input will become associated together in the long-term memory and eventually consolidate into chunks. Upon subsequent encounter, units are not semantically broken down into their component parts, but are typically perceived and processed as a single unit (Ellis 2001: 41).

Importantly, since collocations are not normally taught due to their sheer number in a language, they are subject to more implicit or incidental learning mechanisms, i.e. without learner's conscious attention. An association between items is simply formed through exposure to recurrent patterns and becomes stronger the more often an association is made. Ellis (2001: 41), however, also emphasizes the role of meaning in this process. He maintains that in order for implicit learning of collocations to take place, it is, first of all, necessary that the items are understood and also meaningful to the learner. Moreover, Ellis (2005) maintains that while building of strong association links between word strings might be a long and slow process, conscious attention can cause instant learning. This can probably explain why repeated exposure to collocations does not always lead to their retention in the case of L2 learner. However, as Ellis (2005) claims, once a conscious association has been made, the pattern in question becomes subject to implicit learning and hence to frequency effects. It has been pointed out that a determining factor in usage-based or chunking learning is the frequency of experiencing language instances occurring together (Ellis 2003). This is why

exposure to a language plays a central role in associative or statistical learning that takes place in L1.

When it comes to second/foreign language acquisition, it is still not entirely clear whether a holistic, as described by Ellis (2001), or a more analytical approach (the open choice principle) to language learning is adopted. Due to the dearth of empirical evidence on the subject, Ellis (2003: 72–74) speculates that a similar idiom-based learning pattern can be applied to adult L2 learners as well. Some strong arguments, however, have been put forward against this position. Perhaps one of the clearest opposing views has been expressed by Wray (2002) who, after conducting an extensive literature review on the issue, claimed a fundamental difference between how child L1 and adult L2 learners approach (collocation) learning. Wray argues that a major difference between the first and second/foreign language acquirers lies in the fact that adult non-native learners tend to overlook meaningful connections between words encountered in the input, primarily concentrating on and remembering words in isolation. For example, coming across a collocation such as *make a complaint*, first language acquirers typically process it as a sequence of words and will retain it as a phrasal expression to refer to an act of formal complaining, whereas adult learners will semantically decompose it into word meanings ‘make’ and ‘a complaint’ and store the items as two separate words without any accompanying information about them actually occurring together. Eventually, when an appropriate occasion arrives in the future, non-natives will most likely have no recollection of the standard collocation *make a complaint* encountered before, but rather consider any combination of words with a suitable meaning as equally possible.

Wray explains that this difference in acquiring collocations is largely determined by both social and cognitive factors. At the social level, unlike native speakers, adult learners, especially in a classroom setting, largely lack the pressing need to communicate, which normally pushes L1 learners into memorizing useful communicative strings. Moreover, the acquisition of L2 combinations may even be discouraged by the fact that most social pressure actually comes from the L1 rather than the L2 speech community. This condition, as Wray maintains, is further exacerbated by traditional classroom teaching practices that tend to focus on introducing individual words and grammar as a basis of progress in the target language. Language resources from which learners acquire L2 in the classroom may also substantially differ from the naturalistic environment input. From the cognitive point of view, it is maintained that mature and, typically, literate adult language learners are to a great extent influenced by their analytical faculties urging them to segment sequences into their constituent elements (Wray 2002: 205-6).

Furthermore, Wray (2002) maintains that while negative social effects predominantly concern learners in the classroom context, significantly lacking in necessary socio-interactive pressures, cognitive factors generally apply to both EFL and ESL learners. Although Wray entertains the possibility for L2 learners to be able to achieve a high level of collocation knowledge at some stage, she believes that it is attained in a characteristically different way than in a first language, and hence is inevitably different in kind. Which is why Wray concludes that, because of the influence exerted on L2 acquisition by the social and cognitive factors described, adult second language classrooms should not draw on the principles of child L1 acquisition by exposing learners to the extensive target language input in the hope this will lead to collocation acquisition. She argues that the same goal (e.g. English phraseology) needs to be achieved via different routes, probably by learning collocations consciously.

The current state of research into the subject, however, makes a critical evaluation of the two models difficult. The studies that have empirically put the second model of acquisition to the test are scant, one of these is Durrant & Schmitt (2010) who, contrary to Wray's argument, found that L2 learners do show sensitivity to what items go together in their language input and acquire collocations to which they are frequently exposed. The study indicates that collocational information seems to be retained implicitly, i.e. independent of intentional study or learning techniques. The researchers thus concluded that possible flaws in non-natives' ability to chunk recurrent L2 strings of words may not be a result of a fundamentally different approach to learning, but should be attributed to a lack of sufficient exposure to L2 input (Durrant & Schmitt 2010:179). The findings of this study are congruent with those reported in an earlier work by Durrant & Schmitt (2009), showing non-natives to be quite successful at acquiring many of high-frequency target language combinations. The relatively minor presence of low-frequency collocations in learner data observed was explained by the fact that such combinations are likely to be acquired later than other, more frequent, formulas. This acquisition pattern thus provides evidence in support of the usage-based model of second/foreign language learning, where more exposure to lexical units leads to better acquisition. Durrant & Schmitt explain that a general feeling that non-native language seems unidiomatic is not caused by the fact that it lacks formulaic expressions, but rather native-like diversity, as L2 learners tend to overuse very common collocations and underuse low-frequency, but very salient L1 combinations. Also, a more recent study by Deogratias (2011), showing learners to perform significantly better on high-frequency combinations, provides further support of the idiom-based model of L2 collocation acquisition.

2.4. 2 Factors effecting acquisition: interpreting previous findings

As the above discussion indicates, L2 learners, just like first language acquirers, are in general capable of implicitly learning collocations to which they are frequently exposed. At the same time, as previous research on L2 collocation use demonstrates, chunking process is to a large extent unpredictable, takes a long time and does not always have positive effects. Studies are consistent in finding L2 language to be riddled with deviant and erroneous forms that are highly resistant to change over time or through increased L2 proficiency. In terms of comprehension and processing efficiency, learners again seem to be far from native-like proficiency, showing an overall weak feeling of what words customarily occur together or what combinations are more common than others. It thus appears that collocation learning in the non-naturalistic environment is a challenging task with many open questions and variables that may play different parts in this process. The present sub-section is an attempt to provide an account of major factors that might influence non-native learning of L2 multi-words, drawing on a large body of research on L2 collocation use and knowledge.

First of all, the inherent difficulty of particular collocations can be said to largely determine their “learning burden” (Nation 2001: 23-4), i.e. how difficult they are for the learner to master. Following Nation’s (2001: 36) learning burden principle as applied to collocations, a lexical item has a high learning burden if its semantics, form and usage patterns are not “readily predictable from previous first or second language knowledge”. As a large number of studies demonstrate, the verb component in verb-related combinations appears to be most difficult for the learner as it is probably subject to considerable variation across languages. Particularly hard to acquire were found to be restricted multi-verb units that involve very common verbs, especially in delexical uses (e.g. Bahns & Eldaw 1993; Chi et al. 1994; Howarth 1996; Kaszubski 2000; Nesselhauf 2005). A significantly high learning burden was attested with non-congruent L2 collocations (Gitsaki 1999; Altenberg & Granger 2001; Eftekhari & Rahimi 2011; Yamashita & Jian 2010). An example of this case would be again the combination *make a complaint* which is expressed in the Russian language as *подавать жалобу* [podavat zhalobu] and directly translated as ‘*submit a complaint*’. In this case, there is no way a Russian learner would know how to correctly render the concept referred to by the combination in English unless it is encountered before and learned this way.

Furthermore, previous research has consistently shown that L1 had a strong impact on the amount and type of learners’ collocation errors that also did not seem to reduce as a function of intensive L2 input or increased general L2 competence (e.g. Al-Zahrani 1998; Biskup 1992; Farghal & Obiedat 1995; Howarth 1996; Laufer & Waldman 2011; Nesselhauf

2005). This widespread finding indicates that L1 collocational patterns may strongly interfere with the acquisition of L2 collocations, causing the production of atypical forms. This, in turn, emphasizes the need for taking into account the role of L1 when researching L2 collocation acquisition as it might influence the degree of learning burden of some items depending on their L1-L2 congruency. In the case of the current study, in particular, phrasal verbs can be considered particularly difficult, because very often the meaning expressed by an English verb-noun pair is conventionally rendered in Russian/Ukrainian by means of a one-word item, as in *make excuses* vs. Russian *оправдываться* [opravdivatsia]. Thus, the Russian language can be said to influence learner production in that many multi-word verbs, in addition to resultant erroneous combinations, are also likely to be translated with L1 equivalent one-word items.

Another strong factor in the learning of collocations is the frequency of items attested in the empirical studies reported earlier which found mastery of collocations to be closely related to their frequency, and that less frequent collocations might cause more difficulty for the learner (Deogratias 2011; Durrant 2008; Durrant & Schmitt 2009, 2010; Siyanova & Schmitt 2008). This learning pattern, as mentioned in the preceding discussion, is consistent with the usage-based model of acquisition advocated by Ellis (2001), as well as the connectionist paradigm. A factor strongly linked to the frequency-based learning is exposure to L2 input, in that if an item is frequent it means that it occurs often in the input, and so rich exposure to L2 will lead to retention of frequent forms. The importance of constant and extensive exposure to L2 lexical patterns for their acquisition is widely acknowledged in the literature (Coady & Huckin 1997; Nation 2001; Schmitt 2010) as well as empirically attested (Deogratias 2011; Durrant 2008; Durrant & Schmitt 2009, 2010; Siyanova & Schmitt 2008).

Nevertheless, as a number of studies showed, the acquisition and development of high-learning burden collocations can still remain resistant to chunking via extended L2 input. In particular, Yamashita & Jian (2010) found that non-congruent collocations are not readily acquired even with extensive L2 exposure. Lowie & Verspoor (2004), who investigated the role of L1 similarity and the frequency of items in acquiring English prepositions, also concluded that non-congruent forms would require more time and exposure than congruent ones to enter the new L2 lexicon and hence free active use. These findings thus suggest that mere exposure to (high-learning burden) L2 forms might not be sufficient for their chunking and retention. As Gass (1999) rightly points out, incidental learning of lexical items is most likely to take place if the words are congruent, frequently encountered and known by the learner; otherwise, more intentional learning techniques might be required.

Indeed, empirical evidence suggests that any type of activity that enhances student's engagement with a lexical item/collocation might significantly facilitate its learning (Li & Schmitt 2009). In an overview of instructed L2 vocabulary acquisition, Schmitt (2010: 28) proposed the term 'engagement' to capture a range of factors that have been found to boost vocabulary learning, namely exposure, attention, noticing, manipulation, or time spent on learning lexical items. There is also evidence that the process of being tested on lexical items as well as the number of interim tests can be highly retention beneficial (Schmitt 2010). It is thus argued that, in principle, anything that leads to more engagement might considerably assist acquisition. When applied to the domain of collocations, in particular, this can mean that general involvement with the L2, including more targeted learning of combinations might be necessary to enhance L2 collocation competence.

Mostly on the basis of the results from learner collocation production and reception studies, researchers generally conclude that poor collocation knowledge is a result of the fact that learners are basically unaware of L2 collocations as they are not made salient to them (Gitsaki 1999; Bahns & Eldaw 1993; Farghal & Obiedat 1995; Granger 1998; Hesselhauf 2005). Thus, it seems that first of all raising learners' awareness of collocational associations between words and lexical restrictions might be necessary to improve collocation learning gains and, consequently, production. In fact, there is currently empirical evidence showing that vocabulary/collocation acquisition can be largely facilitated by instruction using explicit or focus-on-form techniques (Laufer & Girsai 2008; Zaferanieh & Behrooznia 2011). In particular, the effectiveness of explicit instruction can be explained by the Noticing Hypothesis proposed by Schmidt (1990), which holds that for acquisition to take place linguistic features need to be first noticed in the input. This is also consistent with Ellis' (2005) view of the importance of conscious attention to forms for their retention.

Specifically, the hypothesis claims that for items to be acquired, learners need to be made aware of them and pay conscious attention to them when encountered in the input. In the case of collocations, explicit instruction may be considered particularly beneficial given the nature of this type of formulaic language. That is, most collocations are semantically transparent as they are usually made up of common individual words as *make a mistake*, *pay a visit*, *give a lecture*, and thus rarely pose comprehension problems. Due to this feature, collocations, when met in the input, are often overlooked or disregarded by both learners and teachers as lacking significance (i.e. importance for understanding). Also, since the use of non-standard collocations, though making sound odd, does not lead to communication disruption, learners' satisfaction with being able to make themselves understood might not contribute to the development of collocation knowledge as well.

Taken together, the points discussed above, suggest that it is not only the amount of L2 input that is decisive, but first of all the quality and type of exposure to input. As a matter of fact, it has been argued in the literature, that a determining factor in acquisition may not be so much exposure per se as the kind of high-quality exposure that is more likely to occur in a socioculturally-integrated environment (Durow 2004 as cited in Schmitt 2010: 140). As with other aspects of L2, a major problem with learning collocations might therefore be a lack of intensive naturalistic exposure to the saturated target language input. Whereas natives come to master collocations in a natural way, that is, subconsciously and gradually via intensive and rich exposure as well as output possibilities as they grow up in their speech community, L2 learners do not have such an opportunity, which thus appears to have an adverse effect on their knowledge of L2 phraseology.

Hence, it seems that with collocations it is crucial how they are learned, which is reminiscent of Wray's (2002) position on collocation acquisition. This argument can be further corroborated by studies that found significant differences in collocation knowledge between learners' across various learning environments, namely naturalistic, ESL and EFL, that apparently all differ not only in the amount of exposure, but also in the ways collocations are acquired/taught there. A case in point is Biskup's (1992) study which revealed that (German and Polish) learners' collocation production differed significantly depending on the emphasis in EFL instruction they received in Germany and Poland. Also Alskaran (2011) and Yamashita & Jian (2010) found ESL users to significantly outperform learners from a foreign language environment. Similarly, Siyanova & Schmitt (2008) discovered that EFL learners with extended stays in a native-speaker environment, i.e. with assumingly rich L2 input and output opportunities, showed better collocation performance than those without this privilege. Moreover, Aghbar's (1990) study showed differences in collocation mastery not only between ESL and native speakers, but also between two naturalistic environments (university teachers of English and undergraduate NS students) .

These findings thus logically lead to proposing a further factor in L2 collocation acquisition, namely the quality of exposure to the L2 input, which will be used here to cover such aspects as the intensity, richness and authenticity of L2 input as well as the type of exposure to the input. The type of L2 exposure, in particular, concerns teaching practices and/or pedagogical focus linked to a particular learning environment (EFL curriculum) in which acquisition takes place. On the other hand, the quantity of L2 input will be used to only refer to the length (total years) of L2 exposure.

2.4.3 Learning in the EFL classroom

The above account clearly indicates that there can be substantial differences between learning environments in which non-natives come to acquire collocations, thus emphasizing the importance of the factor of a learning setting in investigating L2 acquisition of collocations. Considering the focus of this paper on the EFL learner, it is important to describe the learning setting of EFL to understand how it might influence target language knowledge as opposed to other contexts of learning. The present sub-section is thus an attempt to account for learning of collocations as it takes place in the environment of the EFL classroom.

Despite the generally accepted view that psycholinguistic processes underlying learning of a second/foreign language are largely independent of where or how learning occurs, there is certainly no reason to exclude the possibility of differences between learning situations. In fact, there is ample empirical evidence suggesting that the learning environment might substantially influence the process of (collocation) learning (Alsakran 2011; Deogratias 2011; Eftekhari & Rahimi 2011; Gass & Selinker 2008; Yamashita & Jian 2010). Perhaps the most obvious difference concerns the quantity and quality of the target language input available for acquisition in a particular learning setting (Gass & Selinker 2008). It is thus worth taking a closer look at what characterizes the learning context of instructed foreign language (IFL) along these dimensions to understand how it might affect collocation acquisition.

First of all, the IFL environment can be described as significantly constrained in terms of how much and what kind of L2 input there is available for non-native learners, which is in line with Wray's (2002) position. Unlike naturalistic or ESL acquirers, foreign language learners can be said to receive only limited input from which they learn to understand the workings and idiomaticity of the L2. In an EFL environment, the only L2 language non-natives are exposed to is very often limited to the one provided in the classroom. There are thus three major sources of input for IFL learners to draw on: 1) teacher 2) materials, and 3) other learners. The quality of sources 2 and 3 can be, however, significantly undermined through other learners' restricted knowledge of L2, as well as teacher's modifications (linguistic simplifications) often made to the language directed at learners, considered necessary to facilitate comprehension and classroom interaction (Gass & Selinker 2008).

A further limitation of the IFL setting arises from the fact that some language features (e.g. collocations) as well as particular types of situational contexts may not, or very scarcely, be present in the classroom due to obvious time and setting constraints (Spada & Lightbown 2000). This means that some highly recurrent and important collocations related to a particular topic, for instance economics or politics, might never/rarely appear in the classroom

that is inevitably a small-scale imitation of how the target language is acquired in a naturalistic environment. Moreover, opportunities for interaction and practicing essential for developing productive L2 skills can also be seen as fairly limited in a foreign language setting. In light of the above, it can be assumed that the biggest problem with learning collocations in the classroom might be first and foremost lack of sufficient and high-quality exposure to the target language.

Furthermore, in addition to the limitations concerning the intensity and richness of L2 input, collocation acquisition in the IFL setting can be also substantially influenced by a certain type of instruction or emphasis in the curriculum. Research showed, for example, that some instruction types can be especially benefiting for learning collocations in overcoming the effects of frequency and congruency (Gass & Selinker 2008; Spada & Lightbown 2000; Laufer & Girsai 2008). On the other hand, collocation acquisition may be impeded by the classroom's adherence to some favored teaching practices as, for instance, in primarily communicative and content-based classrooms that focus on input-based learning and refrain from using form-focus techniques, thus prioritizing comprehension and fluency skills over lexical accuracy (Laufer & Waldman 2011; Spada & Lightbown 2000).

Lastly, the IFL environment can be strongly affected by a particular attitude to foreign language acquisition in general that might govern the entire teaching process. Specifically, despite the recognized importance of L2 phraseology, a common observation in the context of IFL learning has been that the collocational dimension of L2 is largely neglected, which inevitably results in learners' poor awareness and hence knowledge of collocations. It is commonly true that in the traditional classroom learners are not typically made aware of the syntagmatic associations between L2 words and collocations are rarely included in the foreign language curriculum of most education systems (Fatemeh & Arshad 2011; Nesselhauf 2005). A plausible reason for such neglect is that, in comparison to other linguistic aspects, collocation knowledge is not regarded as problematic for the learner as meriting special pedagogical attention (Biskup 1992; Zaferanieh & Behrooznia 2011) and hence assumed to develop as a mere by-product of general language acquisition. Moreover, Farghal & Obiedat (1995) observed that non-native idiomatic incompetence is the result of instruction that approaches vocabulary as primarily individual lexical items, or slot fillers in the syntactic grid, thus encouraging atomistic processing of vocabulary, inevitably leading to erroneous production of L2 phraseology.

To sum up, instructed EFL environment is characterized, on the one hand, by the limited in quantity and quality L2 input and, on the other hand, affected by traditional teaching practices that tend to neglect the collocational aspect of the L2, focusing either on

grammar, words in isolation or communicative skills at the expense of lexical accuracy and appropriacy. It can thus be concluded that EFL learner's difficulty in mastering collocations might be to a large extent a result of the inherent nature of the instructed EFL environment, rather than a fundamentally different approach to learning L2 as postulated by Wray (2002).

2.5 Chapter summary

The purpose of this chapter was to provide a theoretical background to the present investigation of L2 acquisition of collocations in the EFL setting. It started with a discussion of the phenomenon of collocation by first outlining major paradigms within which collocation was investigated, and then describing the notion as it is used in this thesis, including the definition of the construct of collocation knowledge. Next, the role of collocations in language learning was addressed, followed by a review of previous research on L2 collocation knowledge and use that was organized around two major types of methodology adopted in the field: corpus-based investigations and elicitation studies. The main finding, as pertaining to the research focus of the current study, is that the use of non-congruent verb-noun collocations involving delexicalized verbs is particularly problematic for non-natives. L2 learners seem to experience difficulty with this collocation type even at advanced levels and errors (mostly L1-induced) in their production do not easily reduce as a function of improved L2 proficiency or extensive L2 exposure.

The final sub-section of this chapter was dedicated to the issue of how L2 learners acquire target language collocations. It opened with a description of a general theoretical framework of (L2) collocation learning and proceeded to the discussion of the matter in light of previous L2 collocation research, in particular as regards factors that might influence the acquisition process. The rest of the sub-section dealt with the nature of L2 collocation acquisition taking place in the EFL classroom. The most essential points that can be sifted out from the general discussion in this part of the chapter are the following. Firstly, empirical evidence was provided in support of the idiom-based acquisition of collocations by L2 learners. That is, non-natives are considered capable of implicitly acquiring a large number of collocations to which they are regularly exposed. This implies that any deficiency in the L2 phraseological competence is most likely a result of insufficient exposure to rich and high-quality L2 input. Secondly, the learning environment was shown to have an impact on L2 collocation acquisition, specifically in terms of the quantity and quality of L2 input, where quality refers to both authenticity and intensity of input, as well as to particular teaching methods and practices guided by emphasis in the EFL curriculum. The last essential point is

that inadequate collocation competence by EFL learners could be to a large extent explained by the inherent limitations of the learning context of the EFL classroom. The following section outlines the methodology of the present analysis intended to explore collocation learning in the EFL environment as well as some of the important factors in this process.

3. Research methodology

This chapter aims to describe the methodology adopted in the present study by first outlining the setting in which the study took place, then describing the participants, the material and test administration, followed by the data analysis procedure and the summary of the hypotheses put forward at the beginning of the thesis.

3.1 Setting

The present study was conducted in the context of EFL, including an analysis of two micro EFL learning environments as represented by two different English study programs, in two departments and universities in two different cities in Ukraine. For confidentiality reasons, the exact names of the institutions are not disclosed here. One group of the participants consists of EFL learners enrolled in the teacher-training program in the English department at a state Cherkassy university (hereafter ‘Group A’). The second group includes translation/interpreting students at a private Kiev university in the department of Languages and Literature (hereafter ‘Group B’). Since the universities are part of the same education system of Ukraine, the two learning settings are expected to concur on two major points, namely the national English syllabus approved for higher education institutions by the Ministry of Education in Ukraine, and a general approach to foreign language teaching, e.g. in prioritizing grammar skills as well as rote memorization techniques.

The two learning settings, however, can differ as regards the following aspects: 1) emphasis in the curriculum of the respective study program and consequently teaching practices, e.g. the focus on communicative skills or translation strategies and techniques, 2) hours of instruction, and 3) quality of training and L2 input/output, including teaching materials and text books. Thus, the two EFL learning environments are assumed to vary in the quantity and quality of EFL input, where the quality also includes the type of L2 exposure related to a particular teaching methodology adopted in a respective classroom. Table 2 below provides more detailed background information about the settings obtained through an informal oral interview of the head teachers of the respective departments as well as of some of the students.

As can be seen, the two learning settings indeed differ in both the amount and quality of L2 input. The Cherkassy university English curriculum includes almost twice as many hours of English instruction as that of the Kiev program. Language of instruction in the Cherkassy EFL setting seems to be predominantly English (80%), while in the Kiev university it is split

between English and Russian/Ukrainian. The learning contexts also vary in the emphasis of their study programs and, consequently, also teaching methodologies. That is, they have two different tasks: to prepare specialist in English translation/interpreting and in English teaching, which are most probably achieved by different pedagogical approaches. With regard to English phraseology, as the interview revealed, the curriculum of the two study programs does not explicitly deal with collocations, based on which it is assumed that recurrent combinations are the subject of more implicit learning.

Table 2. **The description of two EFL learning environments**

Area	Cherkassy University	Kiev University
Department	English Studies	Languages and Literature Studies
Program	English Teacher Education	Translation/Interpreting
Hours of English	about 7 hours a week 360 hours a year	about 5 hours a week 200 hours a year
English-related subjects	Phonetics, Grammar, Oral Practice of English, English Teaching Pedagogy	Grammar, Theory/Practice of Translation/Interpreting, Communication Practice
Interim tests	N=16 per semester	N= 15 per semester
Language of instruction	about 80% in English 20% in Russian/Ukrainian	about 50% in English 50% in Russian/Ukrainian
Emphasis in instruction	grammar and lexis, oral skills, language pedagogy, rote learning	grammar and lexis, oral skills, translation/interpreting, rote learning

3.2 Participants

The participants of the study were in total 80 EFL students: 40 from Group A and 40 from Group B, each comprising 20 students from the first and 20 from the last academic year. The participants were all majoring in English, but, as indicated above, in two different study programs: students in Group A were in the teacher-training program and in Group B, studying translation/interpreting, and their age was assumed to range between 17 and 22. Most of the participants speak Ukrainian and to a lesser degree Russian as their mother tongues. Drawing on the information provided in a specially designed questionnaire, the participants started learning English either in grade 1 (at the age of 6) or grade 5 (at the age of 10) in Ukrainian secondary schools at a rate of 3-5 hours a week, which means they had had from about 7 to 11 years of formal English instruction before entering university. Hence, there is a four-year

difference in the L2 exposure length between the ‘early’ and ‘late’ EFL starters. A possible explanation of such a discrepancy can be the fact that in Ukraine, based on the old education system, English as a subject is introduced at national secondary schools in different years, usually in the first or the fifth grade, occasionally in year 3 or 7. Meanwhile, there is a growing tendency to teach English starting from grade 1, which can probably explain a larger proportion of early L2 starters within the Level 1 group than in the Level 4 group (see Table 5.1: Ch.4.4).

All participants had to pass school-leaving as well as university entrance exams on English and their proficiency can be estimated to be at least at a pre-intermediate level for the first-year students, and advanced for the last-year students. Approximately 35% of all test-takers had been engaged with English out-of-classroom before or after entering university or both by, for example, taking additional English classes with a tutor or in a language school, or/and simply by doing regular extra reading, listening to the news, watching TV or communicating in the target language. None of the students reported to have spent longer than 2 weeks in an English-speaking environment. That is, the participants’ exposure to L2 had been largely restricted to the classroom where English is taught as a foreign language subject. Therefore, it is assumed that the learners in this analysis share a comparable level of formal education in English. Finally, in addition to the informants of the study, 6 intermediate and 8 advanced students learning English in different programs participated in the test-development phase. All participants were unpaid volunteers.

3.3 Test instrument and procedure

3.3.1 Test format

The first step in the study was to develop a measurement instrument to elicit the participants’ knowledge of collocations. Given the present focus on the productive side of collocation competence, a special productive collocation test was designed consisting of two parts: a cloze or fill-in-the-blank test (hereafter ‘Test 1’) and a translation task (hereafter ‘Test 2’). A complete test instrument is included in Appendix II. Test 1 and Test 2 aimed to tap into two aspects of productive collocation knowledge, i.e. cued production and free production, respectively. The decision to use two kinds of productive tests was based on evidence of the incremental nature of learning collocations suggesting degrees of knowing collocations at the productive level as well (Li & Schmitt 2009). The intention of using Test 1 and Test 2 was thus to conduct a fine-grained analysis of the learners’ productive collocation mastery.

The adoption of the productive test format for investigating non-native collocations, in general, can be regarded as a justified approach because L2 production is acknowledged to be the clearest evidence of what learners have acquired (Ellis & Barkhuizen 2005: 21). Moreover, the use of controlled production is indispensable when investigating a particular type of lexical items, since unguided production might result in wildly varying combinations (Siyanova & Schmitt 2008) rendering an intended analysis impossible. As regards the type of tests in particular, the adoption of a translation and a cloze test is warranted by the fact that they are two widely-used and established measures of productive collocation knowledge to date (e.g. Al-Zahrani 1998; Biskup 1992; Bahns & Eldaw 1993; Gitsaki 1999). Before administering the test, it was trialled on a small group of intermediate and advanced Russian/Ukrainian students learning English for different purposes, i.e. to become English language or literature teachers or English interpreters/translators. On the whole, the pilot test showed the ability of the test instrument to distinguish between the levels and learning environments, ranking advanced students with intense and frequent exposure to English at the top.

The employed test format is, however, not without its limitations, due to which results should be interpreted with caution to avoid overgeneralizations. First of all, the data elicited experimentally can be seen as inherently flawed in that it basically provides information of a learner's careful style, which is the opposite of naturally occurring data highly preferred in EFL/ESL research (Ellis & Barkhuizen 2005: 48). In other words, the data obtained from the tests used in the study can only show what students can produce under a particular condition (e.g. of taking a test), but it may not predict with certainty what learners can do under more natural conditions of L2 use. With regard to a translation test, a further limitation concerns the fact that L1-L2 translation tends to encourage strong L1 interference (Ellis & Barkhuizen 2005: 38; Farghal & Obiedat 1995), which may negatively influence the production of non-congruent collocations. However, and as was confirmed by the pilot test with some highly proficient EFL speakers, if an item is sufficiently mastered and active in use, its translation as well as cued production will most certainly be correct. That is, only inadequately learned collocations are likely to result in erroneous forms or non-production on the collocation test.

3.3.2 Material

A total of 40 verb-noun collocations were included in the collocation test, equally divided between Test 1 and Test 2. Researchers generally agree that the number of target collocations used in a study should preferably not be too large (over 50) or too small (10-15)

(Barfield & Gyllstadt 2009; Gyllstad 2007). Therefore, 40 items can be considered satisfactory for the present investigation to enable valid conclusions of the results.

For selecting target combinations a number of general criteria had to be met. Firstly, English collocations had to be restricted in Howarth's (1998a) sense and also non-congruent for the Ukrainian/Russian learner. Taking into account the mother tongue of the majority, the test was constructed in the Ukrainian language. It should be noted that Ukrainian can also be effortlessly understood by learners with Russian L1, as the two languages are very similar in grammar and vocabulary as well as in phraseology. The focus on non-equivalent L1-L2 collocations is justified by ample evidence showing this type of combinations to be particularly problematic for learners (e.g. Eftekhari & Rahimi 2011; Gitsaki 1999; Yamashita & Jian 2010; Nesselhauf 2005). Secondly, the verb component of the target collocations had to be frequent, polysemous and/or in a delexicalized meaning, which is widely reported in the literature as causing significant difficulty to non-natives (e.g. Altenberg & Granger 2001; Eftekhari & Rahimi 2011; Kaszubski 2000). Thirdly, the noun constituent had to be a common word, that is, fall within the range of the first 3,000 most frequent word families, which corresponds with the level of average English high-school textbooks in Ukraine. In doing so, it was made sure that individual items involved in the target collocations were known by both the first-year and the last-year students and their misunderstanding did not interfere with the collocation performance on the test. Lastly, the target combinations had to be of different frequencies to enable the investigation of frequency effects.

In particular, the item selection process was carried out following a two-phase procedure. In the first phase, 13 common polysemous English verbs were selected drawing on previous research, including the BNC corpus to establish the range of the item meanings. The resulting list included the following verbs: *break, carry, catch, cut, do, draw, get, give, hold, make, pay, raise* and *take*. For the noun part, again, the BNC was consulted, using the *collocates* function, to extract a number of noun candidates for the 13 verbs. When searching the corpus, only constituents that occurred close to one another (allowing one space in between, e.g. for an article) were considered. The second phase involved manual selection of nouns from the extracted set of options to be combined with the 13 verbs to form target collocations. The selection was guided by the following criteria: the noun had to 1) be common as attested by the *Web VocabProfile/BNC-20* (Cobb 2006), 2) belong to different frequencies (high/low frequency), and 3) form English collocations with no direct Ukrainian/Russian translation equivalents as established by major Ukrainian-/Russian-English bilingual dictionaries (ABBY Lingvo 2008; CEUD 2012).

As a result, a list of 40 combinations that most closely met the criteria was compiled. Additionally, drawing on the BNC frequency information, the target collocations were split into two frequency bands: a) collocations with a BNC frequency of over 100 (hereafter ‘high-frequency collocations’) and b) collocations with a BNC frequency below 100 (hereafter ‘low-frequency collocations’). When determining collocational frequencies in the corpus, maximum one space was allowed separating the verb and the noun part of a target combination in order to accommodate a possible insertion of, for instance, an article or a personal pronoun as in *do **an** experiment* or *break **one’s** promise*. The 40 target collocations were then equally and in a random order divided into two sets to be included in Test 1 and Test 2. The collocations in the two tests were also matched for frequency. A complete list of the 40 target collocations used in the present analysis, including their frequency band information, is shown in Table 3 below.

Table 3. **Target verb-noun collocations and their frequency**

	High-frequency collocations (BNC frequency >100)	Low-frequency collocations (BNC frequency <100)
1	give rise	cut prices
2	make sense	make excuses
3	give birth	get the opportunity
4	make a decision	draw the curtains
5	raise money	give pleasure
6	take action	pay respects
7	make a mistake	give a lecture
8	pay attention	make a complaint
9	make a profit	carry the burden
10	raise a question	draw a conclusion
11	do (no) harm	make threats
12	take a risk	hold a conference
13	get the impression	carry a gun
14	catch fire	catch a cold
15	make a contribution	cut production
16	raise awareness	do a course
17	raise voice	break a promise
18	hold office	break the cycle
19	pay a visit	do an experiment
20	give an example	pay compliments

The next step in the procedure of developing the measurement tool involved designing a two-part collocation test, consisted of Test 1 (a cloze test) and Test 2 (a translation task), each targeting 20 collocations. As mentioned earlier, Test 1 was intended to measure cued production and Test 2 to elicit free production of collocations. For the cloze test, short independent sentences were constructed, each containing a gap to elicit the missing verb pair of a verb-noun collocation in the infinitive form. To limit the scope of verb options, for most test items the first letter of the deleted verb was provided, including the Ukrainian translation equivalent of a target collocation given in brackets at the end of each sentence. In four sentences that involved either very short or assumingly easy verb constituents (e.g. *do a course*), only the L1 translation without the onset letter of the prompted verb was included. To enable contextual authenticity the BNC corpus was employed to sample sentences that would most adequately capture the meaning expressed by target collocations. In this way, the most representative options were selected and/or modified, when necessary, to make sure the described situation was clear and explicit without unnecessary details. Care was also taken to keep the vocabulary involved in Test 1 predominantly neutral and within the range of the first 3,000 most frequent word families.

In Test 2, the other set of 20 target collocations were included in the Ukrainian language to be translated into their English collocation counterparts. The most appropriate Ukrainian translations of these items were selected drawing on some acknowledged Ukrainian-English bilingual dictionaries. Importantly, only L1 equivalents that could be translated unambiguously in both directions (i.e. L1-L2 and L2-L1) were included in the test. Subsequently, to establish the validity of Test 1 and Test 2 as well as the time needed to complete them, the test instrument, as already indicated above, was piloted on a small group of Ukrainian/Russian speaking learners of EFL. As a result, the test had to undergo some slight adjustments. Specifically, in Test 1, two context sentences were reformulated to more accurately capture the meaning of the given collocations. In Test 2, several Ukrainian translations were changed into more appropriate ones to avoid ambiguity.

In addition to the collocation test, a small questionnaire, included at the beginning of the test sheet, was designed to obtain data on the participants' formal and informal English language learning experience, namely the onset of learning English and out-of-classroom L2 engagement. Also, to gain some background information about the English curriculum and teaching practices in the two learning environments under focus (see Table 1 above), an informal oral interview was conducted with each of the head teachers of the respective university department, including some of the students. In this way, the final material of the study included one productive collocation test of two parts, a cloze test and a translation task,

as well as a small written questionnaire on the participants' English learning background (see Appendix II), and an oral interview of the departments' faculty members.

3.3.3 Test administration

After having designed and validated the test instrument, the author contacted a faculty member at the English departments of the universities in Cherkassy and Kiev and asked for permission to administer the collocation test and to conduct an oral background information interview. The purpose of the test and research project was explained and the test material demonstrated in person to responsible parties, upon which permission was granted and an appointment set for a convenient day during a usual 80-minute university session. Subsequently, the collocation test was administered to approximately 80 students (in 4 classes) during their regularly scheduled instruction times. All the subjects of each group/sub-group were tested on one day and all students belonging to a particular group were tested together. The average number of students in a Ukrainian university class amounts to 25. However, since on the day of test administration several students were either absent or did not participate in the study, the number of the test-takers was slightly reduced so that there is an equal proportion of the students in Group A (N=20+20) in relation to Group B (N=20+20), i.e. 20 participants included in the study were from the first academic year and 20 from the last year at each university.

Before the test started, the students were told that the purpose of the study was to investigate non-native learning of English vocabulary. They were assured that the test was anonymous and would not have any influence on their course grade. All test-related details and procedures were carefully explained to the students both orally and in a written instruction included in the test paper. In addition to that, the participants were referred to a sample item at the beginning of the test to ensure the familiarity with the new format. The informants were specifically told not to use dictionaries or consult their neighbor's sheet, but rather leave blank any item they were not sure of. The students were allowed about 30 to 35 minutes to complete the test. After administering the collocation test, the researcher then conducted a brief interview with the head teachers of the respective university departments as well as with a few students to gain as much background information as possible on the amount and type of English instruction in a given study program. This information supported the discussion of the results of the analysis exploring the influence of the quantity and quality of exposure to EFL input on L2 collocation acquisition.

3.3.4 Data analysis and scoring procedure

After collecting the completed tests, every paper was assigned a separate number corresponding to one participant including an indication of the student's belonging to a particular group (A or B) and sub-group (year 1 or year 4). Next, the data obtained from the tests was coded and scored according to a following set of criteria. All responses on the collocation test were scored as either correct or incorrect, where 1 point was awarded for each correct answer and 0 point was assigned to each incorrect answer or unanswered item. The total score for the test instrument was 40, 20 for Test 1 and 20 for Test 2. Spelling mistakes (e.g. **rais a question*) and morphological errors, such as the use of the verb tense deviant from the intended one (e.g. *people got the impression*, instead of *get*) were disregarded. In those cases, where students provided a collocation that did not match the target one, but was nevertheless an acceptable English combination, the item was marked as correct and assigned to a respective frequency band as attested by the BNC. To provide an example of this case, in the translation task, the Ukrainian equivalent of the English target collocation *get an opportunity* was sometimes translated as *get a chance*, which is a synonymous and also an appropriate English combination and therefore judged as correct. For determining appropriacy of such non-target combinations, in addition to the BNC, two well-known collocation reference sources were consulted, the BBI Dictionary of English Word Combinations (BBI; Benson, Benson & Ilson 1997) and the Oxford Collocations Dictionary (OCD 2002). For the analysis of collocation frequency effects, the data was also coded in terms of the participants' scores on high-frequent and low-frequent collocations.

The data was analyzed using the Statistical Package for the Social Sciences (SPSS) 15.05, in particular, descriptive statistics and inferential statistics. The conventional alpha level of 0.05 was set to determine the statistical significance of the results for all the analyses performed. There were five dependent and four independent variables in the study. The dependent variables involved the participants' scores on the entire collocation test and separately on Test 1 and Test 2, as well as on high-frequency and low-frequency collocations. The independent variables were the participants' EFL learning setting, academic level, length of L2 exposure, and the intensity of exposure to EFL, labeled as either 'extra' or 'no extra' EFL engagement. Extra EFL engagement was marked positively when a participant reported in the background questionnaire that s/he was either in an English-focused curriculum class or involved in extra English learning before and/or after entering the university for at least a year (e.g. private classes) in addition to frequent out-of-class L2 engagement that included at least one of the these types of activity: listening to/watching English, talking in or reading English.

Following this, two types of analyses were conducted. Firstly, performance on the collocation test was compared between learners at two proficiency levels, i.e. 40 students in the first academic year (hereafter ‘Level 1’) and 40 in the last academic year (hereafter ‘Level 4’). In particular, the learner data was analyzed in terms of the following aspects: 1) differences in scores on the entire collocation test as well as on the cued vs. free collocation production test; 2) differences in the production of collocations depending on their frequency; 3) the effects of the length of exposure to EFL and the quality of L2 input (extra EFL engagement) on collocation mastery. Secondly, in order to examine the influence of the type of exposure to EFL (the micro EFL learning context), learners representing two different EFL settings, i.e. Group A and Group B, were compared with regard to their collocation performance on the entire test, separately on Test 1 and Test 2, and on collocations from two different frequency bands.

3.4 Research hypotheses

Before embarking on the empirical analysis, I would like to provide an overview of the research hypotheses formulated at the beginning of the thesis.

- 1) Last-year EFL students demonstrate distinctly better productive collocation knowledge than students in their first academic year.
- 2) Participants’ knowledge of collocations at the free production level lags behind their cued productive collocation competence.
- 3) Learners’ performance on high-frequent collocations is better than on low-frequent combinations that are more difficult to master.
- 4) There is no (or a very weak) relationship between participants’ length of EFL exposure and their command of collocations.
- 5) Participants with extra-curricular EFL engagement show better collocation performance than those, whose L2 exposure is limited to the classroom EFL input.
- 6) Learners from two different ‘EFL setting’ groups will demonstrate different collocation performance in that one group will achieve better scores than the other.

4. Results and discussion

This chapter presents the findings of the current analysis, including a discussion. The aim of this study was to explore the acquisition of the productive knowledge of incongruent delexical verb-noun collocations in the context of English as a foreign language. For this purpose, a productive collocation test, consisting of a 20-item cued production test and a 20-item translation task, was administered to a group of 80 Ukrainian/Russian speaking EFL learners in the first and the last academic year as well as from two different EFL study programs. The chapter is organized around the sets of data aimed to address each of the six hypotheses outlined earlier.

4.1 EFL learners' productive collocation knowledge

To test RQ1, which explores learners' productive collocation knowledge in the first vs. last academic year, a comparison of their total-test scores was conducted using an independent samples t-test. Table 4.1 illustrates the descriptive data for this analysis, showing score means, including their percentages and standard deviations. Not surprisingly, as shown in the table, the seniors performed better on the test than the first-year students. The difference between the two proficiency groups in their collocation mastery was also found to be statistically significant ($t(78)=-22,4$, $p<.001$, sig.2-tailed=.000).

Table 4.1. **Score means for the collocation test for Level 1 and Level 4**

Test	Learner level	Mean	%	Std. deviation
Total test (N=40)	freshmen (n=40)	15,70*	39%	3.36
	seniors (n=40)	29,67*	74%	2,05

* $p<.005$

The boxplot in Figure 1 below also demonstrates the percentage of correct answers provided in the test at two levels, where the top whiskers refer to the maximum score achieved by the test-takers (Level 1=25 and Level 4=34) and the lower whiskers refer to the least achieved score (Level 1=10 and Level 4=26). The upper quartile indicates that 75% of the scores fall below the top of the box and above the median, represented by a black line inside each box. The lower quartile points to 25% of the scores that are below the median. Thus, it can be observed that the seniors outperformed the freshmen by the median being higher in the box related to the Level 4 participants than in that representing Level 1.

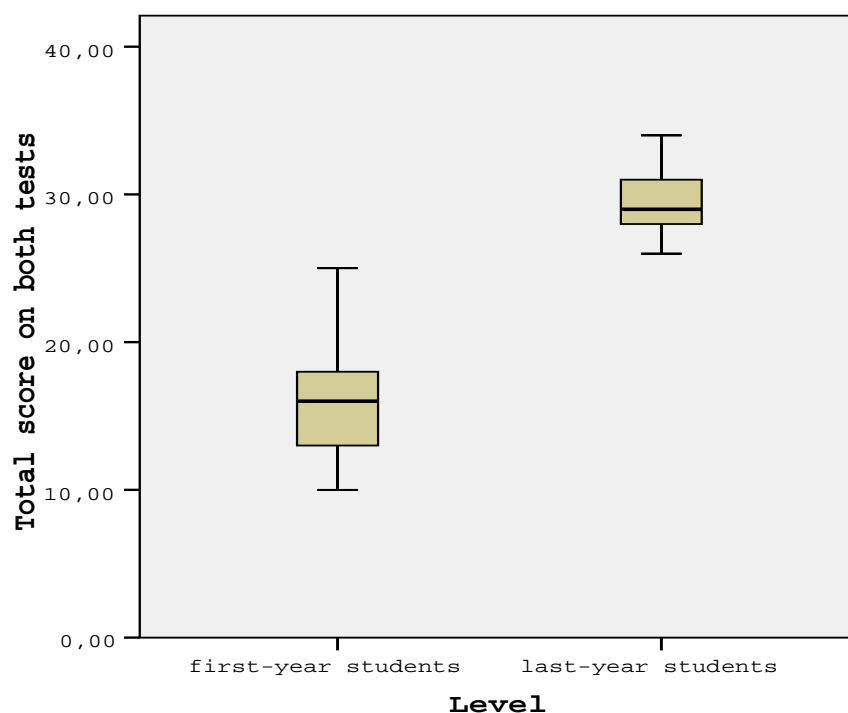


Figure 1. Percentage of correct answers in the test at two levels

The results of this analysis can thus be said to confirm the first hypothesis postulating the Level 4 learners to outperform the first-term students in their productive collocation competence. These results also corroborate the findings of some previous studies that reported collocation knowledge to improve parallel to general L2 proficiency (e.g. Bonk 2001; Deogratias 2011; Gitsaki 1999). At the same time, however, as is also obvious from the data, despite the apparent difference in collocation competence between the two proficiency groups, the learners can be said to demonstrate overall insufficient knowledge of the target collocations both at Level 1, with 16 (out of 40) correct items, and Level 4, with 29 correct items. In percentage terms, the learners acquired nearly 74% of the verb-noun collocations by the end of their studies in relation to the freshmen's knowledge of nearly 40% of the combinations. Thus, there is approximately a 35% increase in collocation competence from Level 1 to Level 4. This is, however, rather moderate given that the knowledge of almost half (40%) of the collocations was already acquired in the first year of study, but did not reach beyond the 80% growth point by the last academic year.

Clearly, to corroborate these findings a longitudinal study involving the same learner population, preferably at three levels (in years 1, 3, 5), and whose L2 proficiency level is formally determined, might be necessary. Nevertheless, the data indicate that the use of the target verb-noun collocations poses a serious challenge to learners. Moreover, an (assumed) increase in the students' general L2 competence (i.e. progress from the first to the last

academic year) does not seem to substantially contribute to their productive mastery of collocations. In this regard, the findings are in line with other studies that found the use of collocations to cause difficulty to learners irrespective of their level of L2 proficiency (e.g. Howarth 1996; Laufer & Waldman 2011; Nesselhauf 2005).

There are two possible explanations for the observed slow expansion of the learners' L2 collocation knowledge. First of all, the non-native repertoire (variety) of L2 collocations can be to a large extent restricted and 'non-inclusive' due to the phenomenon of lexical "teddy bears" (Hasselgren 1994) that holds that learners tend to stick to a limited set of collocations with which they are most comfortable, e.g. L1 congruent, learned well and frequently used or simply favorite combinations (e.g. Aghbar 1990; Durrant & Schmitt 2009; Granger 1998; Hasselgren 1994; Kaszubski 2000). In other words, the students failed to produce a particular number of standard L2 collocations probably because these items were beyond their lexical 'teddy bears' knowledge that also does not appear to significantly expand as a function of extended L2 instruction (from Level 1 to Level 4). Another reason can be fossilization process, as reported in Eftekhari & Rahimi (2011), who found that errors of collocations not properly acquired initially tend to fossilize at an intermediate level, showing high resistance to change as learners progress to higher levels in L2 proficiency. That is, an observed high proportion of the students' atypical combinations represent collocations that were not learned properly at some stage in L2 acquisition process and which with time became deeply entrenched in the interlanguage (fossilized), resulting in the persistence of erroneous forms irrespective of the amount of L2 input or instruction received.

In fact, both the persistence of collocation errors and learners' adherence to a limited set of favored combinations could be induced by the high learning burden of the combinations under study, namely non-congruent delexical verb-noun collocations. This type of multi-words is also reported by other studies to pose considerable challenge to the L2 learner independent of the proficiency level and received L2 input (Altenberg & Granger 2001; Chi et al. 1994; Gitsaki 1999; Kaszubski 2000; Yamashita & Jian 2010). In other words, it means that the target collocations can be particularly difficult to acquire and gain a productive control of, which is why learners prefer to adhere to a conservative repertoire of familiar (easier) and favorite combinations used as 'safe bets', in Granger's (1998:148) words, and, when required to produce forms beyond their knowledge (as in a collocation test), resort to different strategies and draw on available L2 linguistic resources, often resulting in errors.

A discussion of the participants' performance on the test will help support this point as well as provide some examples of the EFL learner specific use of the target collocations. Although no quantitative analysis was conducted of the collocation production, two major

error types could be easily observed in the test-takers' deviant uses of the collocations: 1) L1-induced errors, and 2) overextensions of verb uses (e.g. *do, get, have, make*), followed by 3) a mixture of blends, word-for-word translations or paraphrasing often using one-word L1 translation equivalent. Examples of these categories are:

- 1) **close the curtains* (**закрити** занавіски [zakryty znavisky]), **make experiments* (**робити** експерименти [robyty experymenty]), **pass a course* (**проходити** курс [prohodyty kurs]) instead of the target collocation *do/take a course*, **bring pleasure* (**приносити** задоволення [prynosyty sadovolennia]);
- 2) **do the burden*, **do birth*, **do a mistake*, **get a cold*, **get a rise*, **have harm*, **have office*, **have a risk*, **make harm*, **make conclusions*, **make a conference*;
- 3) **bear the weapons*, **gain the impression*, **elevate your voice*, **pick up a question*, *find excuses/excuse oneself* (for *make excuses*), *to please/*give satisfaction* (for *give pleasure*) *to experiment* (for *do an experiment*).

These instances thus constitute several key strategies learners employ to deal with combinations that are probably unknown or not properly mastered. Similar errors are reported to be frequently produced by L2 learners in other studies (e.g. Biskup 1992; Chi et al. 1994; Farghal & Obiedat 1995; Howarth 1996; Nesselhauf 2005). The default, or a highly preferred, strategy to tackle problematic collocations is obviously L1 transfer, which is not surprising as it was also found to be the case in a large number of previous L2 collocation studies (e.g. Altenberg & Granger 2001; Biskup 1992; Farghal & Obiedat 1995; Hasselgren 1994; Nesselhauf 2005; Yamashita & Jian 2010; Laufer & Waldman 2011).

A special case of deviant collocation use observed in the learner data was an avoidance of a delexicalized verb collocate by means of substitution through an acceptable, but less frequent alternative element, e.g. *reduce/shrink production* (for *cut production*), *cause harm* (for *do harm*), *perform/conduct experiment* (for *do experiments*), *provide example* (for *give an example*), *come to a conclusion* (for *draw a conclusion*). Given that these uses have a lower (BNC) frequency than the target forms, it can be speculated that they are a result of rote learning techniques, often practiced in the Ukrainian education system, rather than a product of frequent exposure to them in the L2 input. In particular, such uses may indicate that some frequent L2 combinations, as noted by Spada & Lightbown (2000), can be of rare occurrence in the EFL input and hence not available for frequency-based learning (chunking). At the same time, a preference for a more formal variant of an L2 collocation, which is less common and thus unlikely to have been implicitly learned before its frequent counterpart, can suggest learner's careful style in producing L2 forms induced by the elicitation technique employed in the study. This means that instead of producing a structure that is an informal but high-

frequency item and thus expected to come first to mind as in a naturally occurring speech, the students apparently made a conscious effort to provide an answer option that they believed was the best for the test, i.e. a less frequent item but of formal, 'correct' style.

With regard to the participants' overall performance, an interesting observation was that collocation items were rarely left unanswered in the test. In other words, most students supplied their answers on nearly every translation or blank-sentence item, whether it was wrong or right. The exception was the first-term students, who often provided either a correct answer or no answer at all. This apparent overconfidence in approaching L2 phraseology may be explained, first of all, by learners' general low awareness of the L2 syntagmatic relationships and lexical restrictedness, which echoes other researchers' observations (Chi et al. 1994; Eftekhari & Rahimi 2011; Granger 1998). As a matter of fact, such a 'never-give up' strategy can be a product of a general attitude or approach to EFL teaching/learning characteristic of the learners' immediate classroom or the education system in Ukraine in general. Yet, it is also possible to assume that in this behavior pattern some other, more psychological, factors might be in play.

Although not directly related to RQ1, an interesting pattern revealed in the descriptive statistics of the analysis is worthy of mention as a last point here. As shown in Table 4.1, and as is also evident in Figure 1, there is a comparatively high standard deviation value in the data of the Level 1 students ($SD=3,36$), showing their scores to be more spread out or heterogeneous than in the case of the fourth-year students. Specifically, some freshmen scored as high as 25, others 16, while other learners achieved only a 10-point score on the entire test, which thus points to a quite uneven level of collocation knowledge among learners of this group. This is probably due to the fact that, before entering the university, the Level 1 learners received L2 instruction in different EFL settings (primary/secondary schools) so that their collocation learning gains might have been influenced in different ways. On the other hand, the last-year students, stemming from the instructional setting (the EFL curriculum) to which they had been exposed for a long period, appear to be rather homogeneous in their mastery of collocations ($SD=2,05$). In particular, this finding emphasizes the importance of the (immediate) EFL learning environment of the learners in collocation acquisition.

To recapture the findings of this analysis, the participants of the study demonstrated overall insufficient productive knowledge of the target collocations at both Level 1 and Level 4, indicating that collocation development does not seem to be parallel to the acquisition of many other aspects of L2 competence. The difficulty in producing collocations was attributed to the inherent nature of the target combinations as well as the limitations of the EFL classroom.

4.2 Degrees in productive knowledge of collocations

To address RQ2, investigating the development of collocation mastery along the cued-free production continuum, the participants' scores on Test 1 (cued production) and Test 2 (free production) were measured and compared by running two paired samples t-tests for each learner level group. The results show that the learners at both levels did significantly better on Test 1 ($t(78) = -23,2$, $p < .001$, sig.2-tailed = .000) than Test 2 ($t(78) = -18,25$, $p < .001$, sig.2-tailed = .000). As descriptive statistics in Table 4.2 illustrate, the Level 1 students' score mean is 6,95 (35%) in Test 1 vs. 8,77 (43%) in Test 2, thus showing the free production of collocations being 10% lower than the cued production. Similarly, with the Level 4 students, the mean score of 12,85 (64%) on Test 2 is lower than on Test 1 with the mean of 16,8 (84%), however, at a slightly higher level of 20%.

The results thus indicate that the learners' collocation competence is better at the cued productive level than the free collocation production, which confirms the second hypothesis of the study stating that the participants' free productive mastery of collocations lags behind the cued productive level. These results corresponds with the findings in Laufer & Girsai's (2008) study, showing learners' performance on the vocabulary and collocation test to be significantly better on the passive recall than on the active recall test. Thus, it appears that free production of collocations is more challenging for learners and might also take longer to achieve than the less demanding cued productive skill. In general, these findings can be said to provide further evidence in support of the incremental nature of collocation learning documented in the studies of Li & Schmitt (2009) and Laufer (1998), which was previously postulated only with relation to individual words (Schmitt 2010).

Table 4.2. **Score means for Test 1 and Test 2 for Level 1 and Level 4**

Test type	Learner level	Mean	%	Std. deviation
Test 1 (n=20)	freshmen (n=20)	8,77*	43,5%	1,79
	seniors (n=20)	16,80*	84%	1,24
Test 2 (n=20)	freshmen (n=20)	6,95*	35%	1,72
	seniors (n=20)	12,85*	64%	1,09

* $p < .005$

As regards the development of the productive collocation competence from the first to the last year of EFL instruction an interesting trend can be identified. As mentioned above, there is a difference in the improvement rate between the productive aspects of collocation

mastery with a larger gap separating the cued production and the free use of collocations among the seniors (20%) than the freshmen (10%). This can indicate that learners' free productive knowledge of collocations does not develop in parallel with the cued productive mastery. In other words, the knowledge of collocations at the free productive level seems to reach a plateau at some point of L2 learning, from which a gap starts to grow separating this aspect and the cued productive knowledge as learners move to higher (academic) levels of L2 proficiency. However, to support this finding a comparative analysis of the collocation performance by learners across three university years (1, 3 and 5) would be necessary. One obvious reason for the differences between the two productive skill aspects might be that the knowledge of some collocations acquired at the cued productive level apparently does not transfer into their free active use. This, in fact, can also be a result of the earlier mentioned lexical "teddy bears" and fossilization phenomena.

The asymmetry in the development of L2 collocation competence is an attested fact in the literature. Notably, Laufer (1998), who investigated the relationship between different aspects of vocabulary mastery, reported the gap between learners' receptive and productive knowledge to increase at advanced levels. Similarly, Farghal & Obiedant (1995), who also used a close test and a translation task to study L2 collocation knowledge, found that learners performed considerably worse on L1-L2 collocation translation, which was apparently more difficult than prompted collocation production.

On the whole, on the basis of the results of this analysis it can be speculated that the observed slow and uneven increase rate of the participants' collocation knowledge was brought about by the high learning burden of the target collocations, as noted earlier. At the same time, the poor collocation performance, especially on the free production task, can suggest that EFL classrooms generally lack activities directed at developing free active productive skills of collocations necessary to push passive knowledge into active production. A follow up study comparing non-native receptive and productive knowledge of the target collocations, however, will be needed to tell more about the exact relationship between the two aspects and their development.

4.3 Frequency of collocations and their mastery

To address RQ3 the target collocations were selected along the variable of high-low frequency. Starting with the assumption that frequency is an important determinant of collocation acquisition, the participants were expected to score better on high-frequency collocations than on low-frequency combinations. In order to test this, two paired samples t-

tests were run for each level, using the number of correct answers for the high-frequency (n=20) and low-frequency (n=20) target collocates. The results showed statistically significant differences between the answers for the collocations of the two frequency bands at both Level 1 ($t(39)=12,41$, $p<.001$, sig.2-tailed=.000) and Level 4 ($t(39)=16.80$, $p<.001$, sig.2-tailed=.000). Descriptive statistics for this analysis are presented in Table 4.3, showing mean scores, including their percentages and standard deviations.

As can be seen in the table, an overall better performance on frequent collocations is obvious, with the participants at both Level 1 and Level 4 scoring 13 % and 18% percent, respectively, better on the high-frequent combinations ($M= 9,17$ and $M=16,65$) than on the low-frequency items ($M=6,52$ and $M=13$). The increase in knowledge of frequent collocations (of 38%) from Level 1 to Level 4 was also 5% higher than for the low-frequency items with 33% of growth rate. The standard deviations show the scores for the both types of collocations to be quite homogeneous and evenly distributed, again except for the freshmen groups ($SD =1,97$; $1,63$) demonstrating a slight diversity in their responses.

Table 4.3. **Learners’ performance on frequent and low-frequent collocations**

Collocation type	Learner level	Mean	%	Std. deviation
HifrqColl	freshmen (n=20)	9,17*	45%	1,97
	seniors (n=20)	16,65*	83%	1,36
LowfrqColl	freshmen (n=20)	6,52*	32,5%	1,63
	seniors (n=20)	13,05*	65%	1,13

Note. HifrqColl= high-frequent collocations; LofrqColl =low-frequent collocations
* $p<.005$

The results thus clearly demonstrate a difference in the L2 performance between the two types of collocations in favor of the high-frequency ones at both academic levels, which, in turn, emphasizes the importance of frequency in collocation acquisition process. This finding therefore confirms the third hypothesis stating that learners perform better on frequent collocations than on low-frequency combinations that are more difficult to master. The results of this analysis are not surprising as the positive influence of the frequency of collocations on their learning is widely documented in the literature (Deogratias 2011; Durrant 2008; Durrant & Schmitt 2009; Siyanova & Schmitt 2008). Moreover, this finding provides further evidence in support of the usage-based model of L2 acquisition (Ellis 2001; Tomasello 2003), which suggests L2 learners’ ability to acquire collocations frequently met in the L2 input.

A closer look at the learner collocation production, however, reveals that performance on some low-frequency items (4 in total) was better than on high-frequency collocations. The same phenomenon was observed in Deogratias' (2011) study that found 24% of low-frequency items that were better identified than more common collocations. In this study, collocations that were readily produced despite their low frequency in English are: *make a complaint*, *make threats*, *do a course*, and *carry the burden*. It can be speculated that the production of the correct combinations with *make* was in most cases purely a result of the overextension strategy rather than actual knowledge. The collocation *do a course* was probably easy for the university students due to its obviously high frequency in the academic-related discourse to which the learners were frequently exposed. Finally, the target item *carry the burden* was rarely an area of difficulty in the translation test possibly because of the similarity to its Ukrainian counterpart *нести тягар* [nesty tyagar] 'wear/carry the burden'.

Combinations that seem to have caused difficulty despite their high frequency include: *do harm*, *take action*, *raise money*, and *raise awareness* that proved to be particularly problematic in the test. It is hard to say with certainty what could be a source of difficulty in the case of these items as there might be several reasons, from confusion brought about by some words' broad semantic field (e.g. *give pleasure*=*satisfaction/enjoyment/fun*), through a lack of familiarity with the item constituents (e.g. *raise*, *awareness*) to misunderstanding caused by ambiguity in the Ukrainian translation of the combinations provided to elicit the forms. Nevertheless, such instances also serve to show that frequent exposure to (problematic) collocations may not be the be-all and end-all of their acquisition in the EFL setting and that there are some other, probably stronger, factors to take into consideration, in particular the quantity and quality of exposure to L2 input that are addressed in the following section.

4.4 The quantity and quality of EFL input in learning collocations

In exploring research questions 4-6, the assumption was that collocation acquisition can be influenced by the quantity and quality of L2 input available for the learner in a particular EFL learning environment. As described earlier, in the current study, the quantity of EFL input refers to the length of exposure to EFL, while the quality concerns the intensity of L2 exposure as well as the type of EFL engagement (the curriculum of a given study program), including L2 input richness and authenticity, e.g. in terms of language of instruction and teaching/course materials. To examine the factor of the EFL quantity, the collocation performance of learners with differing EFL onset times was tested and compared. In investigating the variable of the EFL quality, learners with different EFL engagement

intensity, as well as from two different EFL settings were compared on their collocation mastery.

Table 5.1 presents data on the participants' English learning background collected via a specially designed questionnaire. In particular, the table includes statistics on years of exposure to EFL and the number of learners involved in extra-classroom engagement with EFL. Based on the EFL onset year (1-3 grade or 5-7 grade), the freshmen were estimated to have 7 and 11 and the seniors 11 and 15 years of EFL exposure, taking into consideration the EFL onset time of the majority of the test-takers, namely either the first or the fifth grade. As shown in Table 5.1, the learner data is very similar in the length of formal education in English, demonstrating an overall 4-year difference between the early and late EFL starters at both levels. Furthermore, it can be seen that there are about twice as many learners (34) with early EFL onset in the Level 1 group than there are among the last-year students (16). Such a high percentage of early starters among the freshmen, as mentioned earlier (Ch.3.2), can reflect a new (approximately 5-year old) trend in the foreign language education policy of Ukraine to introduce English in the school curriculum as early as possible.

With regard to extra EFL engagement, the learners are very similar in both level groups in that nearly one third (35%) of all the participants (on average 15 out of 40) were involved in out-of-classroom EFL learning. In fact, the data in the table confirm the researcher's personal observation made during her studies that the majority of university students tended to be only occasionally involved in extra English practice or learning. The reason for this is mostly either time constraints because of the intensive study schedules, or some students' complacency about their English knowledge, particularly those, who started learning English in the first year of school.

Table 5.1. **Participants' EFL learning background**

Level	Length of EFL exposure				With extra engagement
	11 years	7 years	15 years	11 years	
freshmen (n= 40)	34 (85%)	6 (15%)			15 (37,5%)
seniors (n=40)			16 (40%)	24 (60%)	14 (35%)

4.4.1 Length of EFL exposure

To test RQ4 the performance of the students with the two different EFL onset times (the first vs. the fifth grade) was compared using two independent samples t-tests for each level group. Table 5.2 summarizes key statistical results of the test, including total-test score means for the learners with shorter vs. longer EFL exposure at the two academic levels.

Table 5.2. **Length of EFL exposure and total-test scores for freshmen and seniors**

Level	Length of exposure	Mean	Std. deviation	df	t	Sig. (2-tailed)
freshmen (n=40)	11 years (n=34)	15,6	3,36			
	7 years (n= 6)	16	2,28	9,88	-.318	.757
seniors (n=40)	15 years (n=16)	29,5	1,75			
	11 years (n=24)	29,7	2,26	37,0	-.458	.650

As revealed in the table, both the freshmen and seniors are very similar in their production of the collocations regardless of the length of EFL exposure. There is only a minor difference between the two L2 onset groups at both levels of maximum 1%, surprisingly, in favor of the learners with 4-year shorter exposure to the target language. As standard deviations indicate, there is again an uneven distribution of the data within the Level 1 group. Since the homogeneity of variances was violated, as assessed by the Levene's Test for Equality of Variances ($p=0.20$ for Level 1 and $p=0.36$ for Level 4), the data in the bottom row of the inferential statistics output was used, which showed all paired differences to be statistically insignificant ($p>.001$, sig.2-tailed=.757 for Level 1, and $p>.001$, sig.2-tailed=.650 for Level 4). From this data it can be inferred that 4-year longer exposure to EFL does not significantly contribute to better collocation performance, which thus confirms the fourth hypothesis postulating no positive (a very weak) relationship between the length of EFL exposure and the learners' collocation mastery.

Overall, the results of this analysis suggest that in the case of high-learning burden collocations, mere exposure to L2 input or, more specifically, 4-year long exposure to L2 in the EFL classroom is insufficient to bring about considerable collocation learning gains. This finding can be supported by Yamashita & Jian's (2010) and Laufer & Waldman's (2011) studies that found non-congruent verb-noun collocations to be hard to acquire even with extensive exposure to the target language. Also, the studies that focused on restricted verb-

noun collocations, including common delexical verbs, found such collocations to be highly resistant to ordinary learning and L2 input (Eftekhari & Rahimi 2011; Howarth 1996). A further support comes from the argument that exposure to the target language input, though essential for collocation acquisition, might not always be as important as its quality (Schmitt 2010: 140; Durrant 2008).

A surprising pattern in the data was that students with shorter EFL exposure showed a slightly better performance (1%) than later EFL starters. This result can be explained in light of the researchers' personal experience concerning general attitudes to EFL learning in her homeland. There is a widely held belief in Ukraine that L2 users who started learning English in the first grade or in the kindergarten have automatically better L2 competence than those, who began with EFL later on. I assume that this particular attitude have an effect on students' behavior regarding English learning practices. In particular, it was observed that early EFL starters were usually quite satisfied with or even complacent about their English knowledge and therefore rarely engaged in L2 learning outside the classroom. On the other hand, late L2 starters tended to be more motivated to learn the language in addition to the classroom activities (especially with a private tutor), probably, due to the feeling that they needed to improve or catch up with their knowledge of the L2 because they started learning it "late".

Thus, it can be assumed that the participants with shorter EFL exposure fared better on the test because they were apparently more intensively involved in English learning/practice than the early starters. This, in turn, lends support to the importance of the factor of the quality of EFL engagement in collocation acquisition explored in the following sub-section. Finally, it should be noted that the results of this analysis need to be treated with caution as they might have been affected by the small and also unevenly distributed number of the learners in the two 'length of EFL exposure' groups at each level (see Table 5.2).

4.4.2 Intensity of EFL exposure

In exploring RQ5, the test performance between the learners with varying intensity of EFL engagement was compared using two independent samples t-tests for each level. Mean scores for the correct answers provided by the participants with or without extra EFL engagement, involving the data for the first-year and last-year level, are illustrated in Table 5.3 below. Since the homogeneity of variances was violated also in this analysis, based on $p=0.40$ in both levels as determined by the Levene's Test for Equality of Variances, the data in the bottom row of the inferential statistics output was treated as valid. The analysis revealed paired differences between the scores of the students with and without extra engagement with

EFL at both academic levels, which were also statistically significant. ($t(23,9)=-5,78$, $p<.001$, sig.2-tailed=.000 and ($t(23,9)=-6,43$, $p<.001$, sig.2-tailed=.000)

Table 5.3. **EFL Engagement and test scores for Level 1 and Level 4**

Level	EFL engagement	Mean	Std. deviation	df	t	Sig. (2-tailed)
freshmen (n=40)	Extra (n=15)	18,7*	2,78	23,95	-5,78	.000
	No extra (n=25)	13,8*	2,14			
seniors (n=40)	Extra (n=14)	31,7*	1,54	23,58	-6,43	.000
	No extra (n= 26)	28,57*	1,33			

Note. Extra= Extra EFL engagement, No extra= no extra EFL engagement.

* $p<.001$

As is evident in Table 5.3, the learners at both levels with extra EFL engagement achieved higher scores on the collocation test ($M=18,7$ and $M=31,7$) than the students involved in no extra EFL activities ($M=13,8$ and $M=28,5$). For Level 1, the difference in the scores between the ‘EFL engagement’ groups is about 12%, while for the senior groups it is slightly smaller with 8%. The data thus indicates better performance on the test by the learners involved in out-of-classroom English learning, which confirms the fourth hypothesis postulating that more intensive involvement with L2 will lead to better production of collocations. Thus, these results provide evidence in support of the importance of engagement as well as the role of L2 quality in vocabulary/collocation acquisition also widely documented in the literature (Durrant 2008; Li & Schmitt 2009; Schmitt 2010; Wray 2002). Furthermore, considering that the variable of the length of EFL exposure was found to have no significant effect on the learners’ collocation performance, it can be concluded that personal extra-classroom engagement with L2 might be a stronger factor in acquisition of high-learning burden collocations than the actual exposure to L2 in the EFL learning setting.

The analysis is, however, not without its limitations that need to be addressed in further research. In particular, the data on learners’ EFL engagement used in the study was exclusively based on the student’s reports about their L2 learning experience provided in the EFL learning background questionnaire. This information was thus not verified or empirically supported, e.g. through experimental evidence. Furthermore, the study would definitely benefit from a larger sample of L2 learners than involved in this analysis. Finally, in exploring effects of learners’ extra-curriculum EFL learning, it seems that some other important factors,

in particular motivation, learning habits and strategies as well as the quality (richness and authenticity) of EFL input, may play an important part in need of further examination.

4.4.3 The EFL classroom

To test RQ6, exploring collocation mastery of students from two EFL learning settings, the score means of the Group A (Cherkassy) freshmen and seniors were compared with their Group B (Kiev) counterparts by running an independent samples t-test for each level. The results for test-section scores and total-test scores for the freshmen are shown in Table 5.4 and for the seniors in Table 5.5 in Group A and Group B, respectively. As can be seen in Table 5.4, the Kiev freshmen scored slightly better both on the entire test ($M=16,6$) and separately on Test 1 ($M=9,35$) and Test 2 ($M=7,25$) than the Cherkassy freshmen ($M=14,8$ and $M=8,20$, $M=6,65$). However, only paired differences in the performance on Test 1 ($M=9,3$ vs. $M=8,20$) were found to be statistically significant ($t(38)=-2,12$, $p<.05$, sig.2-tailed=.04).

Table 5.4. Test score means for freshmen in Group A and Group B

Test type (n= items)	Group	Mean	Std. deviation	df	t	Sig. (2-tailed)																
Test 1 (n=20)	Cherkassy	8,20*	1,43	38	-2,12	.041																
	Kiev	9,35*	1,95				Test 2 (n=20)	Cherkassy	6,65	1,56	38	-1,10	.277	Kiev	7,25	1,86	Total score (N=40)	Cherkassy	14,8	2,78	38	-1,70
Test 2 (n=20)	Cherkassy	6,65	1,56	38	-1,10	.277																
	Kiev	7,25	1,86				Total score (N=40)	Cherkassy	14,8	2,78	38	-1,70	.091	Kiev	16,6	3,70						
Total score (N=40)	Cherkassy	14,8	2,78	38	-1,70	.091																
	Kiev	16,6	3,70																			

p* <0.05

As regards the last-year learner level, as shown in Table 5.5 below, the Kiev seniors again produced a slightly higher proportion of correct collocations on the entire test than their Cherkassy counterparts ($M=30$ vs. $M=29$) and this difference is statistically significant ($t(38)=-2,17$, $p <.05$, sig.2-tailed= .03). Group B also performed statistically better on Test 2 (free production) ($M=13,2$ vs. $M=12,4$, $t(38)=-2,44$, $p <.05$, sig.2-tailed=.02), whereas no paired differences between the groups in the performance on Test 1 (cued production) were found to be significant. Standard deviations indicate less dispersion in the senior data than among the freshmen in both EFL groups.

Table 5.5. Test score means for seniors in Group A and Group B

Test type (n= items)	Group	Mean	Std. deviation	df	t	Sig. (2-tailed)
Test 1 (n=20)	Cherkassy	16,4	1,14	38	-1,83	.075
	Kiev	17,1	1,26			
Test 2 (n=20)	Cherkassy	12,4*	.88	38	-2,44	.019
	Kiev	13,2*	1,16			
Total score (N=40)	Cherkassy	29.0*	1,77	28	-2,17	.036
	Kiev	30.3*	2,13			

p* < .005

Finally, in order to test the performance on low-frequency vs. high-frequency collocations by the participants from the two EFL settings, only the last-year learner data was used as it contained more responses on the test and therefore deemed to be more suitable for a group comparison analysis. Moreover, since the last-year students share the same EFL learning background, i.e. the curriculum of the teacher training or translation/interpreting program, it is more reasonable to focus on their performance in analyzing effects of the two EFL learning environments in question. For this final analysis, means of the learners' correct responses on the collocations of the two frequency bands were compared between the two 'EFL setting' groups, using an independent samples t-test. The key test results, including descriptive statistics, are illustrated in Table 5.6.

Table 5.6. Scores on high-/low-frequent collocations for Group A and Group B

Item type	EFL group	Mean	Std. deviation	t	Sig. (2-tailed)
HifrqColl	Cherkassy	16,5	1,19	-.688	.595
	Kiev	16,8	1,54		
LowfrqColl	Cherkassy	12,5*	1,05	-3.48	.001
	Kiev	13,6*	.94		

Note. HifrqColl= high-frequent collocations; LofrqColl =low-frequent collocations

*p < .001

As the data in the table demonstrate, the Kiev students (Group B) again achieved slightly higher scores on both types of the collocations ($M=13,6$ vs. $M=12,5$), but the difference between the groups is statistically significant only in the production of the low-frequent collocations ($t(df)=-3.48$, $p < .05$, sig.2-tailed=.001). This finding is consistent with

the results of the earlier conducted analyses, showing the Kiev students to fare better than the Cherkassy group on the entire test as well as the translation task. These results thus confirm the fifth hypothesis of the study stating that learners from different EFL settings differ in their collocation mastery. More specifically, the observed differences in collocation performance between the Kiev and Cherkassy learner groups indicate that the learning environment, be it EFL, ESL or a more immediate one as in this case, is an important factor in collocation learning. The significance of the learning context in which L2 collocation acquisition takes place is also documented in the literature. In particular, Siyanova & Schmitt (2008) found that a year spent in the target language environment substantially contributed to better collocational intuition of EFL learners. The studies of Alsakran (2011) and Yamashita & Jian (2010) also showed that ESL users, enjoying the privilege of learning from the high-quality L2 input, demonstrated more advanced collocation knowledge than learners whose exposure to English was limited to the EFL classroom.

Further evidence in support of the importance of the (immediate) EFL learning environment, although not coming directly from these results, can be found in the distribution of the scores in the freshmen data evident in the descriptive statistics outputs across the study. The conducted analyses consistently revealed standard deviation values to be somewhat higher within this learner group. In particular, this indicates, as noted above (Ch.4.1), that collocation performance at the freshmen level varied considerably from quite high (25 out of 40 score points) to significantly poor (10 out of 40 score points), thus suggesting the influence of the different EFL settings (schools), where the learners received L2 instruction before their training at the respective university, on their knowledge of the collocations.

Now, the important question is what, in terms of the amount and quality of L2 input, could have brought about the observed differences in the collocation performance between the EFL groups. To answer this I analyzed the results in light of the background information describing the curriculum design of the two EFL learning environments, i.e. the Cherkassy (Group A) and the Kiev (Group B) programs. As regards the quantity, or the period of EFL instruction, the two curricular are the same as they are both 4-year long programs. The main point of difference (see Table 2 in Ch.3.1), however, is the quality of EFL input, i.e. the intensity and richness of L2 as well as the emphasis in the respective syllabus, which is reflected in the methodology applied in the classroom, including course materials. Specifically, the Cherkassy classroom provides more intensive EFL instruction (360 hours of EFL a year) than the Kiev setting (200 hours a year), and it is also mainly in English (80% vs. 50% in Kiev). Nevertheless, the Kiev students, despite having received fewer hours of English instruction, showed somewhat better performance on the collocation test than the

Cherkassy group with more intensive EFL curriculum. From this it can be concluded that a major cause of difference in the collocation performance discovered was probably the English teaching methodology employed in the classrooms, namely the type of exposure to EFL.

Furthermore, considering that the emphasis in the Kiev program syllabus is apparently on developing translation/interpreting skills, it can be speculated that the Kiev learners benefited from more form-on-focus teaching methods, i.e. more conscious learning techniques which are at the heart of any translator/interpreter education program. The fact that the Group B students fared better on the translation task can lend further support to this argument. Benefits of focus-on-form (including translation) activities for learning vocabulary and collocations are attested in previous experimental studies. In particular, Laufer & Girsai (2008) found that the contrastive form-focused (L1-L2) analysis and the translation task group considerably outperformed, both receptively and productively, the group who primarily received implicit teaching instruction.

Thus, it can be hypothesized that the type of EFL exposure, or EFL instruction, played a major role in distinguishing between collocation mastery of the two 'EFL setting' learner groups. Undoubtedly, a follow up study may be needed to corroborate the findings as well as to further explore some other important variables in learning collocations in the EFL classroom, e.g. richness and authenticity of L2 input. What is important for the present purposes, however, is that on the basis of the results obtained one might argue that differences between EFL classrooms certainly exist, and most likely in the quality of EFL input, which thus makes this an essential factor when researching L2 collocation acquisition.

5. Conclusions

This section aims to present conclusions drawn from the empirical analysis carried out in the previous section. It starts by pointing out the significance of the study, then introduces the main findings and draws conclusions. After this, pedagogical implications of the study are addressed, followed by a discussion of limitations of the analysis as well as suggestions for further research.

5.1 Significance of the study

Before presenting a summary of the main findings and conclusions, it is worth pointing out some merits of the current study in terms of what makes it distinct from previous research and how it contributes to the existing body of knowledge in the field. Firstly, the research questions addressed in this study concern four dimensions of L2 acquisition of collocations: the role of L2 proficiency, the frequency of collocations, and the role of quantity and quality of L2 input in the EFL learning context. Previous studies have only focused on one or two of the mentioned aspects. What is more, the present investigation analyzed a more nuanced L2 productive collocation competence, distinguishing between cued production (using a cloze test) and free active production (using L1-L2 translation) of collocations, whereas previous research has tended to examine either L2 learners' production or reception of collocations.

Secondly, the study focused on L2 collocation acquisition as it occurs in the EFL environment, also exploring effects of two different EFL learning settings (curricular). Most previous investigations, on the other hand, seem to have ignored the role of the subjects' learning environment, often only distinguishing between native and non-native speakers of English sampled from several different EFL/ESL settings (educational institutions). Thus, studies exploring immediate EFL learning contexts are conspicuously absent in current L2 collocation research.

A third positive distinction of the study is that, in recognizing the importance of learners' L1 in acquiring and using L2 collocations, it concentrated on the analysis of EFL learners sharing the same L1. In previous research, however, the L2 learner data often constituted a pool of EFL/ESL learners from different first language backgrounds. Moreover, the focus on learners with Ukrainian/Russian L1, which were rarely involved in investigations, makes this thesis a particularly valuable contribution to the field in furthering knowledge about collocation acquisition by EFL learners with languages other than so far widely investigated (e.g. Arabic, Chinese, German or Spanish).

5.2 Overview of the main findings and conclusions

The present study intended to investigate the acquisition of non-congruent delexical verb-noun collocations by adult Ukrainian/Russian speaking learners in the EFL environment. Specifically, two main aims guided the study. The first aim was to explore the development of the L2 productive collocation knowledge from the first to the last academic year as well as at the cued and free productive level. The second aim was to examine some external factors that might influence L2 collocation acquisition in the EFL setting, in particular the frequency of collocations and the quantity (length) and quality (intensity and type) of exposure to EFL. To address these aims overall six research questions along with the hypotheses were formulated and subsequently positively answered. In the following, an overview of the main findings of the analysis is provided, followed by a general conclusion.

Firstly, as the total-test results revealed, the EFL learners demonstrated an overall insignificant increase in productive competence of the target collocations from the first to the fourth academic year, with the free productive mastery lagging behind the cued production of collocations, as showed by the scores on the translation and the cloze task. It was thus concluded that knowledge of collocations does not develop in parallel with other aspects of second language proficiency. Moreover, as observed in the learners' collocation performance, the difficulties with the items are most likely linked to the nature of the target collocations, i.e. non-congruent combinations involving a verb in delexical uses, widely attested to have a considerably high learning burden (e.g. Altenberg & Granger 2001; Gitsaki 1999; Yamashita & Jian 2010). A large proportion of L1-based errors in the data, in particular, suggest learners' heavy reliance on L1 lexical patterns when dealing with unknown or poorly mastered collocations.

Secondly, the EFL learners demonstrated an overall better performance on the high-frequency items than on less common collocations that supposedly cause more difficulty. This finding provides further support for the well attested fact in L2 collocation research that the frequency of collocations is closely related to the success with which they are acquired (e.g. Deogratias 2011; Durrant & Schmitt 2009; Nation 1990; Siyanova & Schmitt 2008). In particular, this pattern of acquisition indicates that EFL learners are capable of retaining L2 forms to which they are regularly exposed, which corresponds to the usage-based model of language acquisition (Ellis 2001; Tomasello 2003) and at the same challenges the prevailing view of learners' primarily non-idiomatic approach to L2 processing and learning (Wray 2002).

Thirdly, the quantity (length) of exposure to EFL input showed no significant effects on the learners' collocation competence, whereas extensive engagement with L2, i.e. extra-curricular L2 learning/practice, was found to contribute to better collocation production. This can be taken to mean that it is not so much the exposure to the L2 input that influences collocation acquisition in the EFL environment, but the quality of this exposure, which in turn supports the view expressed earlier about the overriding importance of the quality (intensity and L2 richness) of input in L2 acquisition (Nesselhauf 2005; Schmitt 2010: 140; Siyanova & Schmitt 2008). At the same time, these findings suggest that the target collocations are especially difficult and hence might require a very long time to master properly in the EFL context, which is why the 4-year difference in EFL exposure was probably not enough to distinguish between the participants' collocation performance significantly.

Fourthly, differences were discovered in the collocation production between the Kiev (translators) and Cherkassy (English teachers) learners with the former group slightly outscoring on the entire test as well as on the translation task. It is plausible to assume that the observed better collocation mastery of the Kiev learner group was due to the teaching methodology employed in the Kiev EFL classroom directed at developing translation/interpreting skills that hinged on form-on-focus (explicit learning techniques) methods found to be particularly beneficial for collocation acquisition (Laufer & Girsai 2008; Zaferanieh & Behrooznia 2011). A major implication of this finding is that the EFL learning environment can and should be pedagogically manipulated to the interests and benefits of the learner.

On the basis of the overall results, it can thus be concluded that non-congruent delexical verb-noun collocations present a serious challenge to Ukrainian/Russian L1 learners of EFL, whose implicit acquisition in the inherently limited EFL learning environment, though occurs, is very slow and unstable. This is largely in line with previous studies that reported high-learning burden collocations to constitute a major stumbling block for the L2 learner irrespective of the proficiency level or the amount of L2 input received (Chi et al. 1994; Gitsaki 1999; Howarth 1996; Kaszubski 2000; Granger 1998; Laufer & Waldman 2011; Nesselhauf 2005; Yamashita & Jian 2010). In particular, the participants' observed poor knowledge of the collocations and its insignificant growth from the first to the last academic year indicate learners' overall low awareness of L2 collocations as legitimate and important lexical items, which mirrors observations made by other researchers in the field (Chi et al. 1994; Eftekhari & Rahimi 2011; Granger 1998).

Furthermore, the interplay of three factors was identified to positively influence acquisition of the target collocations in the EFL learning setting: the frequency of

collocations, extra-curricular EFL engagement and the type of exposure to EFL input (instruction), with the overriding importance of the two latter ones (collectively referred to as the EFL quality). Clearly, the exact relationship between these factors, as well as the role of other strong variables, such as motivation and learning strategies, are in need of further investigation. In light of the results produced, however, it can be concluded, that if collocations are to be effectively mastered in the EFL classroom, instead of giving priority to the quantity of L2 instruction, e.g. by trying to introduce English as a foreign language as early as possible, emphasis should be placed on its quality. Specifically, increasing the quality of EFL input would mean combining rich and extensive exposure to L2 (implicit techniques), in order to assist frequency-based learning, with special pedagogical treatment (explicit techniques). Drawing on explicit and implicit methods in teaching is widely documented in the literature as yielding most effective results in L2 vocabulary/collocation acquisition (Durrant & Schmitt 2010; Ellis 2005; Laufer & Girsai 2008; Li & Schmitt 2009; Nation 2001; Schmitt 2010; Yamshita & Jian 2011).

Lastly, the results of this study serve as a warning to L2 collocation researchers to be careful when generalizing about a particular type of L2 learners (e.g. sharing the same L1 and learning environment) from a wide group of learners with varying L1s and L2 learning backgrounds. In their longitudinal study on collocation acquisition by a Chinese-speaking MA student over a period of one academic year, Li & Schmitt (2009) rightly observed that acquisition of formulaic languages can be very idiosyncratic. This is why different variables, particularly those related to the L2 learner profile, need to be carefully controlled for when researching L2 collocation acquisition.

5.3 Pedagogical implications

Perhaps one of the major implications of the findings of the study is that a fundamental change in the mainstream EFL teaching pedagogies in Ukraine needs to be introduced by shifting the focus from individual words to the syntagmatic dimension of the target language. It is important that collocations and formulaic language in general are systematically integrated into the foreign language curriculum and are treated as an essential L2 component on a par with grammar. The observed poor productive collocation mastery of the EFL learners is most likely a result of their generally low awareness of L2 collocations, a large proportion of which remain unnoticed in the input and hence is not retained. This phenomenon can be largely explained by the ‘noticing’ hypothesis (Schmidt 1990), which holds that input does not transform into intake for linguistic items unless they are noticed or consciously attended

to. Along these lines, it appears that only after collocations have made salient, and also acknowledged as legitimate lexical items, both on the part of the teacher and learner, some substantial progress can be expected in collocation acquisition in the EFL classroom.

Furthermore, as demonstrated in the study, a 4-year long period of exposure to EFL input might not be sufficient for the productive knowledge of the target collocations to develop substantially. The important implication is that acquisition of collocations in the EFL context needs to be supported by explicit instruction techniques to override both frequency and item learning burden effects. In practice, of course, it is unrealistic to expect, mostly due to the inherent time constraints of the setting, that all the myriad of collocations, beyond the high-frequency ones (Nation 2001), can be covered in the classroom in any systematic way and at the expense of grammar and/or communicative skills. Therefore, the primary responsibility for this type of learning should lie with EFL learners themselves. The role of language educators in this process, however, is essentially in making collocations salient and important to the learner by means of diverse classroom activities, instruction techniques and learning materials. In other words, the first task is to raise learners' general awareness of collocations in the L2 and possible difficulties they may present (Howarth 1996; Nesselhauf 2005).

At the very basic level, an awareness-raising task could be carried out by providing an explanation of differences between L1-L2 collocations on the basis of a selection of the most representative examples. Additionally, learners can be asked to translate the items into their mother tongue to acquire a better feeling for peculiarities involved in the use of L2 collocations. Following Lewis (1997), various matching activities can also be very useful in drawing learners' attention to the significance of L2 syntagmatic links and lexical constraints. What is important is to present collocations in authentic contexts to expose learners to rich, natural language essential for building native-like lexical associations and expanding vocabulary competence. For this purpose, an excellent tool will be an L2 corpus where multiple authentic sentences can be extracted to present target collocations as well as to form the basis of further learning and practice activities.

In order to develop in learners a habit of thinking of L2 vocabulary in terms of phraseology and lexical patterning rather than individual words, it is important, particularly at advanced levels, to avoid presenting new vocabulary in isolation. It is thus a good strategy to introduce new lexical items with their most frequent collocates embedded in authentic sentences and also to organize a follow-up practice activity around these typical uses. Importantly, pedagogical treatment of L2 collocations should always be provided with reference to learners' first language that, as was shown in this study, is an important factor in

learning L2 phraseology. In this way, it would be possible to predict potential problems certain collocations might present to L2 learners and devise more accurate teaching tools to tackle them.

Moreover, a significant amount of EFL classroom time needs to be dedicated to teaching collocations that are important for use due to their high frequency in the L2 or those that are especially hard to learn, as for example non-congruent delexical verb-noun collocations. In particular, teaching efforts should concentrate on developing collocation mastery at the productive level which was revealed to pose a serious challenge to the EFL learner. This can be done by employing a combination of communicative tasks and focus-on-form activities, as in fill-in-the-blank exercises to supply the missing element of target collocations, translation tasks, retelling or writing stories, composing dialogues or role playing using particular collocations. Especially useful for developing fluency with high-learning burden collocations can be automaticity developing techniques and activities, such as rote learning, repetition and oral drilling (Nesselhauf 2003).

Explicit inductive instruction based on concordance collocation worksheets may also significantly facilitate collocation learning. Following this method, learners are presented with a concordance line containing a lexical item in authentic contexts and asked to work out a typical collocation pattern the word enters and its typical meaning senses. In a follow up activity, for example a fill-in-the-blank exercise may be designed for practicing the collocations in the same/similar set of sentences. Concordance-based activities are regarded as especially useful in dealing with collocations involving common delexical verbs that enter a wide range of lexical patterns, exhibiting varying meanings and uses (Altenberg & Granger 2001).

Furthermore, it is very important to encourage EFL learners to expand their phrasal repertoire beyond lexical “teddy bears” (Hasselgren 1994), which are highly responsible for learner language to sound non-idiomatic and lacking diversity. Different strategies to deal with L2 collocations outside the classroom can be suggested, such as keeping a lexical notebook of collocations learners meet in the input, using collocation cards for systematic review and practice, or noting down from various sources and learning a certain number of new collocations every week. In addition to this, the importance of using collocation dictionaries should be pointed out. What is more, a strong motivating factor in learning collocations both in and outside the classroom for the learner can be regularly conducted various collocation quizzes and tests, as testing is generally acknowledged to positively affect collocation learning and L2 acquisition on the whole (Gyllstad 2008; Schmitt 2010).

Finally, in order to enhance implicit collocation learning, it is crucial to expose EFL learners to a large amount of authentic and rich L2 input, for instance, via extensive reading and listening activities. Durrant & Schmitt (2010), in particular, stress that input-based acquisition can be most effective in the case of activities that enable repeated exposures to the same target collocations within a short period of time. The task of EFL classroom educators is thus to make sure their learners receive intensive and high-quality input to English from a wide range of sources, both within and outside the classroom.

As a closing remark, it should be noted that the suggestions provided here are made primarily with regard to the findings of this study, i.e. focusing on EFL learners sharing the same (Ukrainian/Russian) L1, and therefore may not be entirely applicable in other classroom settings. Moreover, most provided ideas for teaching are only suggestions made drawing on the key principles of successful vocabulary/collocation acquisition attested in the literature, which is why their pedagogical effectiveness is in need of empirical testing. Nevertheless, it is hoped that the present discussion will help raise Ukrainian EFL practitioners' awareness of the importance of learning and teaching collocations, and that the provided suggestions will prove useful in facilitating collocation acquisition in the EFL environment and in Ukraine EFL classrooms in particular.

5.4 Limitations of the study

All in all, it can be said that the present investigation adequately addressed the initially raised research questions and provided valuable insights into the nature of collocation acquisition in the EFL learning environment. The study, however, has a number of limitations and shortcomings, due to which the results obtained are to be interpreted with caution. Firstly, the test material was limited to one type of English collocations, namely verb-noun. Although the use of a large number of combination types can yield data difficult to analyze and draw conclusions from (c.f. Gitsaki 1999), focusing on a single collocation type, on the other hand, may restrict the generalizability of the results to L2 collocation knowledge in general.

Secondly, no in-depth analysis of the participants' collocation production as regards deviant uses, i.e. error types and error sources, was conducted. Although some discussion of the learner performance was provided, it was beyond the scope of this thesis to investigate, categorize and explain in greater detail deviant uses in L2 collocation production. Such analysis is necessary if areas of difficulty related to the use of particular collocations are to be identified for developing special learning/teaching strategies and activities.

Thirdly, L2 proficiency of the participants, ranked to be at the intermediate (in the first year) and advanced (in the last year) level, was not formally determined, but assumed on the basis of the students' level in the education system. This approach was based on evidence from previous research that showed learners at higher levels of study to exhibit advanced L2 collocation knowledge (e.g. Gitsaki 1999; Gyllstad 2007). The problem with this assumption, however, is that it is not self-evident that progression to a higher level in the education system presupposes growth in general L2 proficiency. That is, it is possible that some first-term students possess better L2 knowledge than those in the second or fourth term. Thus, to draw valid conclusions about the relationship between L2 proficiency and collocation knowledge, learners' L2 competence needs to be determined by means of appropriate measurement tools. Moreover, this kind of investigation should be complemented by longitudinal analysis, exploring the development of productive collocation mastery in the same learner population.

Fourthly, in exploring the factor of the quality of EFL input (extra EFL engagement and type of EFL exposure) no data was available on the intensity (e.g. how many hours per week) and the type (personal learning strategies) of out-of-classroom EFL learning, as well as on the exact methodology employed in the two EFL settings under study. This kind of information would be necessary to confirm the findings of the study and also to identify some other strong factors in collocation acquisition in the EFL environment.

A final limitation concerns the fact that the study was based on the data involving a relatively small number of participants, particularly in the case of the between 'EFL setting' groups comparison (20 per group) as well as in the analysis of the effects of the length of EFL exposure and extra EFL engagement, with the smallest learner group consisting of only 6 students. Thus, in order to increase the quality of the study and support the findings it is important to draw on a larger learner sample.

5.5 Suggestions for further research

As the present thesis was designed to be largely exploratory in that it set out to examine the uncharted territory of collocation acquisition in the Ukrainian/Russian EFL classroom, it stands to reason that there are a number of areas and issues in need of further investigation and consideration. In particular, in order to both corroborate the findings of the study and to expand the existing body of knowledge about EFL collocation acquisition, two major research designs can be suggested.

First of all, a study concentrated on the role of the immediate EFL learning environment would be necessary to support the results of the current analysis as well as to provide new

insights into the nature of non-native collocation acquisition in general. Moreover, because factors influencing collocation learning in the two different EFL settings were not systematically addressed in the thesis, it would be interesting to examine such important variables as the quality of EFL input (richness, authenticity) and teaching methodologies employed in the EFL classroom. For testing the effectiveness of different teaching methods (i.e. implicit or more explicit) for collocation learning, an experimental classroom-based study can be considered as most suitable. An inclusion of a learner group from a further EFL setting, as well as an overall large learner sample, would certainly increase the validity of the study. Finally, such a study would probably need to focus on students in their last year of training in a given learning environment as in this case the influence of a target EFL classroom on collocation acquisition is more obvious.

In a second line of research on EFL collocation acquisition, the focus of investigation can be the factors of the quantity (length) and intensity of exposure to EFL as well as the quality (richness, authenticity) of L2 input, including the relationship between these variables. Specifically, to enable a closer examination of the development of collocation knowledge, learners' collocation performance needs to be compared across three academic levels, i.e. the first, the third and the last year of study. Longitudinal research methods can be viewed as particularly beneficial in this case. As for the effects of extra-curricular engagement with EFL, in particular, it would be interesting to examine this variable in relation to the factors of learners' motivation as well as personal learning strategies and techniques. Also, an analysis of L2 collocation competence along the receptive-productive dimension would yield particularly valuable insights into the development of L2 collocation knowledge in the EFL context.

Finally, to increase the generalizability of the study in general, the following methodological improvements can be suggested. Firstly, the test material can be extended by inclusion of two or three more major collocation types, for example adjective-noun and adverb-adjective. Secondly, to avoid ungrounded conclusions about the relationship between learners' L2 competence and collocation knowledge it is necessary to formally determine learners' level of L2 proficiency using an appropriate measure. Moreover, a closer analysis and classification of deviant uses of collocations in learner production would help better understand problems posed by the target collocations, and thus produce more accurate suggestions for their teaching.

References

- ABBY Lingvo Multilingual Dictionary: on compact disc*. 2008. (3rd ed.). <http://www.ABBY.ru>
- Aghbar, Ali-Asghar. 1990. "Fixed expressions in written texts: implications for assessing writing sophistication". In: Paper Presented at a Meeting of the English Association of Pennsylvania State System Universities, October.
- Aitchison, Jean. 2003. *Words in the mind*. (3rd ed.). Oxford: Blackwell.
- Alsakran, Rayed A. 2011. "The productive and receptive knowledge of collocations by advanced Arabic-speaking ESL/EFL learners". MA thesis, Colorado State University.
- Altenberg, Bengt. 1993. "Recurrent verb-complement constructions in the London-Lund Corpus". In Jean, Aarts; Pieter de Haan; Nelleke, Oostdijk (eds.). *English language corpora: design, analysis and exploitation. Papers from thirteenth international conferences on English language research on computerized corpora, Nijmegen 1992*. Amsterdam: Rodopi, 227-245.
- Altenberg, Bengt. 1998. "On the phraseology of spoken English: the evidence of recurrent word-combinations." In Cowie, Anthony P. (ed.). *Phraseology: theory, analysis, and applications*. Oxford: Oxford University Press, 101-122.
- Altenberg, Bengt; Granger, Sylviane. 2001. "The grammatical and lexical patterning of *make* in native and non-native student writing". *Applied Linguistics* 22(2), 173-195.
- Al-Zahrani, Mohammed Said. 1998. "Knowledge of English lexical collocations among male Saudi college students majoring in English at a Saudi University". PhD thesis, Indiana University of Pennsylvania.
- Bahns, Jens; Eldaw, Moira. 1993. "Should we teach EFL students collocations?". *System* 21 (1), 101-114.
- Bardovi-Harlig, Kathleen. 2002. "A new starting point? Investigating formulaic use and input in future expressions". *Studies in Second Language Acquisition* 24(2), 189-198.
- Barfield, Andy; Gyllstad, Henrik (eds). 2009. *Researching collocations in another language*. New York: Pilgrave Macmillan.
- Benson, Morton; Benson, Evelyn; Ilson, Robert. 1997. *The BBI dictionary of English word combinations*. Amsterdam: John Benjamins.
- Biber, Douglas; Johansson, Stig; Leech, Geoffrey; Conrad, Susan; Finegan, Edward. 1999. *Longman grammar of spoken and written English (LEG)* London: Longman.
- Biskup, Danuta. 1992. "L1 influence on learners renderings of English collocations: a Polish/German empirical study". In Arnaud, Pierre J L.; Bejoint Henri (eds.). *Vocabulary and applied linguistics* London: Macmillian, 85-93.
- Bonk, William J. 2001. "Testing ESL learners' knowledge of collocations". In Hudson, Thom; Brown, James Dean (eds.). *A focus on language test development: expanding the*

language proficiency construct across a variety of tests. Honolulu: University of Hawaii Second Language Teaching and Curriculum Center, 113-142.

Carter, Ronald. 1998. *Vocabulary: applied linguistic perspectives* (2nd edition). London: Routledge.

Chi Man-lai, Amy; Wong Piu-yiu, Kitty; Wong Chau-ping, Mabel. 1994. "Collocational problems among ESL learner: a corpus-based study." In Flowerdew Lynne; Tong, Anthony K. (eds.). Hong Kong: University of Science and Technology, 157-165.

Coady, James; Huckin, Thomas (eds.). *Second Language Vocabulary Acquisition*. Cambridge: Cambridge University Press.

Cobb, Tom. 2006. Web VocabProfile BNC-20 v3.2. <http://www.lex Tutor.ca/vp/bnc/>.

Collins English-Ukrainian Dictionary. 2012. New York: HarperCollins Publisher.

Cowie, Anthony Paul (ed.). 1998. *Phraseology: theory, analysis, and application*. Oxford: Oxford University Press.

Durrant, Philip Lee. 2008. "High-frequency collocations and second language learning". PhD thesis. University of Nottingham.

Durrant, Philip Lee; Schmitt, Norbert. 2009. "To what extent do native and non-native writers make use of collocations?". *Language Teaching* 47(2), 157-177.

Durrant, Philip Lee; Schmitt, Norbert . 2010. "Adult retention of collocations from exposure". *Second Language Research* (26)2, 163-188.

Eftekhari, Maryam; Rahimi, Ali. 2011. "The effect of delexicalization of common verbs on the collocational competence of Iranian EFL students". 2011. *Procedia Social and Behavioral Sciences* 15, 3941-3946.

Ellis, Nick C. 2001. "Memory for language". In Robinson, Peter (ed.). *Cognition and second language instruction*. Cambridge: Cambridge University Press.

Ellis, Nick C. 2003." Constructions, chunking, and connectionism: the emergence of second language structure". In Doughty, Catherine J.; Long, Michael H. (eds.). *The handbook of second language acquisition*. Oxford: Blackwell.

Ellis, Nick C. 2005. "At the interface: dynamic interactions of explicit and implicit language knowledge". *Studies in Second Language Acquisition* 27(2), 305-52.

Ellis, Rod; Barkhuizen, Gary. 2005. *Analysing learner language*. Oxford: Oxford University.

Farghal, Mohammed; Obiedat, Hussein. 1995. "Collocations: a neglected variable in EFL". *International Review of Applied Linguistics* 33(4), 315-331.

Fateme, Ebrahimi Bazzaz; Arshad, Abd Samad. 2011. "The use of verb-noun collocations in writing stories among Iranian EFL learners". *English Language Teaching* 4(3), 158-163.

- Firth, John Rupert. 1957. "A synopsis of linguistic theory". In Palmer, Frank Robert (ed). *Selected papers of J.R. Firth 1952-59*. London: Longmans, 168-205.
- Gass, Susan. 1999. "Discussion: incidental vocabulary learning". *Studies in Second Language Acquisition* 21, 319–333.
- Gass, Susan M.; Selinker, Larry. 2008. *Second language acquisition: an introductory course* (3rd ed). New York: Routledge.
- Gitsaki, Christina. 1999. *Second language lexical acquisition: a study of the development of collocational knowledge*. London: International Scholars Publications.
- Granger, Sylviane. 1998. "Prefabricated patterns in advanced EFL writing: collocations and formulae". In Cowie, Anthony Paul (ed.). *Phraseology: theory, analysis, and acquisition*. Oxford: Oxford University Press, 145–160.
- Gyllstad, Henrik. 2007. "Testing English collocations". PhD thesis, Lund University.
- Halliday, Michael A. K. 1966. "Lexis as a linguistic level". In Bazell C. E.; Catford J. C.; Halliday M. A. K.; Robins R. H. (eds.). *In Memory of J.R. Firth*. London: Longman, 148-62.
- Hasselgren, Angela. 1994. "Lexical teddy bears and advanced learners: a study into ways Norwegian students cope with English vocabulary". *International Journal of Applied Linguistics* 4(2), 237-260.
- Herbst, Thomas. 1996. "What are collocations: sandy beaches or false teeth". *English studies* 77 (4), 379-393.
- Hill, Jimmie. 1999. "Collocational competence". *English Teaching Professional* 4(1), 43-50.
- Hoey, Michael. 2005. *Lexical priming: a new theory of words and language*. London: Routledge.
- Howarth, Peter Andrew. 1996. *Phraseology in English academic writing: some implications for language learning and dictionary making*. Lexicographica Series Maior 75. Tuebingen: Max Niemeyer
- Howarth, Peter Andrew. 1998a. "Phraseology and second language proficiency". *Applied Linguistics* 19(1), 24-44.
- Howarth, Peter Andrew. 1998b. "The phraseology of learners' academic writing". In Cowie, Anthony Paul (ed.) *Phraseology: theory, analysis, and application*. Oxford: Oxford University Press, 161-186.
- Kaszubski, Przemek. 2000. "Selected aspects of lexicon, phraseology and style in the writing of Polish advanced learners of English: a contrastive, corpus-based approach". PhD thesis. Adam Mickiewicz University.
- Langacker, Ronald W. 1987. *Foundations of cognitive grammar: theoretical prerequisites*. Stanford: Stanford University Press.

- Laufer, Batia. 1998. "The development of passive and active vocabulary in a second language: same or different"? *Applied Linguistics* 19(1), 255–271.
- Laufer, Batia; Girsai, Nany. 2008. "Form-focused instruction in second language vocabulary learning: a case for contrastive analysis and translation". *Applied Linguists* 29(4), 694-716.
- Laufer, Batia; Waldman, Tina. 2011. "Verb-noun collocations in second language learning writing: a corpus analysis of learners' English". *Language Learning* 61(2), 647-672.
- Lewis, Michael. 1993. *The lexical approach*. Hove: Teacher Training Publications.
- Lewis, Michael. 1997. *Implementing the lexical approach: putting theory into practice*. London: Language Teaching Publication.
- Li, Jie; Schmitt, Norbert. 2009. "The acquisition of lexical phrases in academic writing: a longitudinal case study". *Journal of Second Language Writing* 18, 85-201.
- Lin, Phoebe M.S.; Adolphs, Svenia. 2011. "Sound evidence: phraseological units in spoken corpora". In Barfield, Andy; Gyllstad, Henrik (eds.). *Researching Collocations in Another Language*. New York: Palgrave Macmillan, 49-59.
- Lowie, Wander M.; Verspoor, Marjolijn. 2004. "Input versus transfer? The role of frequency and similarity in the acquisition of L2 prepositions". In Achard, Michel; Niemeier, Susanne (eds.). *Cognitive Linguistics, Second Language Acquisition, and Foreign Language Teaching*. Mouton de Gruyter, Berlin, Germany, 77-94.
- McCarthy, Michael. 1990. *Vocabulary*. Oxford: Oxford University Press.
- Mel'cuk, Igor I. 1998. "Collocations and lexical functions". In Cowie, Anthony Paul (ed.) *Phraseology: theory, analysis, and application*. Oxford: Oxford University Press, 24-53.
- Milton, James. 2009. *Measuring second language acquisition*. Bristol: Multilingual Matters.
- Nation, Paul I. S. 1990. *Teaching and learning vocabulary*. New York: Newbury House.
- Nation, Paul I. S. 2001. *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nattinger, James R.; DeCarrico, Jeanette S. 1992. *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Nesselhauf, Nadia. 2003. "The use of collocations by advanced learners of English and some implications for teaching". *Applied Linguistics* 24(2), 223-242.
- Nesselhauf, Nadja. 2005. *Collocations in a learner corpus: studies in corpus linguistics*. Amsterdam: John Benjamins.
- The Oxford Collocations Dictionary for students of English*. 2002. Oxford: Oxford University Press.

- Pawley, Andrew; Syder, Frances. 1983. "Two puzzles for linguistic theory: native like selection and native like fluency." In Richards Jack C.; Schmidt, Richard (eds.). *Language and communication*. London: Longman, 191-226.
- Peters, Ann M. 1983. *The units of language acquisition*. Cambridge: Cambridge University Press.
- Schmidt, Richard. 1990. "The role of consciousness in second language learning". *Applied Linguistics* 11, 17-45.
- Schmitt, Norbert. 2000. *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schmitt, Norbert. 2010. *Researching vocabulary: a vocabulary research manual*. London: Palgrave Macmillan.
- Schmitt, Norbert; Carter, Ronald. 2004. "Formulaic sequences in action: an introduction". In Schmitt, Norbert (ed.). *Formulaic sequences: acquisition, processing and use*. Amsterdam: John Benjamins, 1-22.
- Sinclair, John. 1987. "Collocation: a progress report". In Steele, Ross; Threadgold, Terry (eds.). *Language topics: essays in honour of Michael Halliday*. Amsterdam: John Benjamins, 319-331.
- Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Siyanova, Anna; Norbert, Schmitt. 2008. "L2 Learner production and processing of collocation: a multi-study perspective." *The Canadian Modern Language Review* 64(3), 429-458.
- Spada, Nina; Lightbown, Patsy M. 2000. *How languages are learned*. Oxford: Oxford University Press.
- Stubbs, Michael. 1995. "Collocations and semantic profiles: on the cause of the trouble with quantitative studies". *Functions of Language* 2 (1), 23-55.
- Stubbs, Michael. 2001. *Words and phrases: corpus studies of lexical semantics*. Oxford: Blackwell.
- Stubbs, Michael. 2004. "A quantitative approach to collocations". In Allerton, David J.; Nesselhauf, Nadia; Skandera, Paul (eds.). *Phraseological units: basic concepts and their application*. Basel: Schwabe, 107-119.
- Tomasello, Michael. 2003. *Constructing a language: a usage-based theory of language acquisition*. Cambridge: Harvard University Press.
- 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.

- Yamashita, Junko; Jian, Nan. 2010. "L1 Influence on the acquisition of L2 collocations: ESL users and EFL learners acquiring English". *TESOL Quarterly* 44 (4), 646-668.
- Zaferanieh, Elahen; Saeedeh, Behrooznia. 2011. "On the impacts of four collocation instructional methods: web-based concordancing vs. traditional method, explicit vs. implicit instruction". *Studies in Literature and Language* 3(3), 120-126.
- Zhang, Xiaolin. 1993. "English collocations and their effects on the writing of native and non-native college freshmen". PhD thesis. Indiana University of Pennsylvania.

Appendix I: A summary of key characteristics of the reviewed studies on L2 collocation knowledge

Study	Method	Subjects	Target collocation types	Findings/conclusions
Aghbar (1990)	cloze test	22 English teachers 44 native undergraduates 97 advanced ESL learners	50 lexical collocations: V+N	-ESL scored worst, performing well only where <i>get</i> was involved, often overusing it
Biskup (1992)	L1>L2 translation	34 L1 German university students of English	(?) lexical collocations V+N, Adj +N	-different strategies used by learners from different EFL settings -strong L1 influence
Bahns & Eldaw (1993)	L1>L2 translation (vocabulary test) cloze test	58 L1 German university students of English in years 1,2, 3	15 lexical collocations V+N	-no relationship between collocational mastery and lexical knowledge - L2 collocations pose a problem
Zhang (1993)	analysis of essays (corpus) cloze test	30 mixed L1s college freshmen 30 native speakers	27 types of lexical and grammatical collocations	-poor NNS writers used fewer lexical collocations -collocations contributes to fluency in writing
Chi et al. (1994)	HKUST L2 learner corpus	L1 Chinese intermediate ESL students	collocations with delexical verbs (<i>do, get, have, make, take</i>) V+N	- great difficulty in matching L2 words appropriately - poor sense of salience of L2 collocations
Hasselgren (1994)	translation of texts	50 Norwegian EFL first-year students of English +60 NSs	lexical collocations Adv+Adj, V+N	- a small number of core verbs (<i>give, get, take, show, have, know, keep, tell make</i>) overused -reliance on familiar/easy items - strong L1 influence
Farghal & Obiedat (1995)	L1>L2 translation cloze test	57 L1 Arabic university students of English	22 common lexical collocations Adj+N, N+N	-overall deficient productive knowledge; reliance on lexical simplification strategies - lack of awareness of collocations - strong L1 influence

				-translation is harder than close test
Howarth (1996)	analysis of L2 essays NS corpus	10 (mixed L1) students of English	lexical type: free combinations, restricted collocations and idioms V+N	- learners underuse restricted collocations - erroneous combinations often as a result of blends - no correlation between collocations knowledge and language proficiency
Al-Zahrani (1998)	cloze test writing section of TOEFL test	81 EFL Saudi students in academic years 1, 2, 3, and 4	50 lexical collocations: V+N	- significant difference in collocations knowledge only between students in 1 and 4 term, - collocations are challenging and their learning takes long time -correlation between collocation knowledge and general proficiency
Granger (1998)	analysis of L2 essays receptive recognition task NS essays	56 L1 French advanced learners of English (+56 English NSs)	lexical collocations: Adv+ Adj	-learners significantly underused target collocations, but over-rely on a small set of combinations - many deviant uses - lack of salience for L2 collocations
Gitsaki (1999)	analysis of essays L1>L2 translation cloze test	275 L1 Greek high-school students (in years 1, 2, and 3)	37 types of non-congruent grammatical and lexical collocations	-collocation knowledge develops in parallel with L2 learners' overall language proficiency -V+N collocations are most difficult
Kaszubski (2000)	L2 learner corpus NS corpus	L1 Polish, Spanish intermediate and L1 Polish, French advanced learners of English	collocations with delexical verbs (<i>be, do, get, have, make, take</i>) V+N	-NNS underuse, but overuse of a small set of collocations - common delexical verb-noun collocations are challenging
Altenberg & Granger (2001)	two L2 learner corpora NS corpus	L1 Swedish, French advanced learners of English	collocations with high-frequency verbs, delexical uses, verb <i>make</i>	- even advanced learners have difficulty with high-frequency verbs like <i>make</i> - delexical uses are particular problematic even at advanced levels
Bonk (2001)	cloze test	98 university students (mixed L1s) of	50-item test	- a close relationship between

	receptive recognition test adapted TOEFL test	different subjects	lexical and grammatical collocations: V + N, V + prep, (fig. use of verb)	collocational knowledge and general English proficiency
Nesselhauf (2005)	analysis of L2 essays (corpus)	207 L1 German university students of English	2,082 (tokens) lexical collocations: V + N	-one third of NNS collocations are unacceptable -verb-noun collocations are most difficult, with the verb part frequently misused - factors of difficulty: L1-L2 congruency, and degree of restriction - length of exposure to EFL classroom had no positive effect, but length of stays in an L2 environment had positive effects
Siyanova & Schmitt (2008)	analysis of L2 essays (Russian L1) vs. NS corpus off-line and on-line receptive recognition task	31 L1 Russian + 9 mixed L1 (mainly Vietnamese, Russian, Chinese) both advanced ESL and EFL learners	31 frequent + 31 infrequent lexical collocations: Adj+N	-NNS were similar to NSs in the use of appropriate collocations -NNS demonstrated poorer intuitions than NSs, and were slower -extended exposure to L2 environment contributes to better collocation intuitions
Yamashita & Jian (2010)	receptive recognition task +recording of reaction time	20 NSs university students 24 L1 Japanese ESL advanced English users 23 L1 Japanese EFL university students	24 congruent +24 incongruent lexical collocations: Adj+N, V+N	-ESL learners significantly outperformed EFL learners both in accuracy and speed -L1 and input influence acquisition - non-congruent collocations are particularly hard to acquire
Alsakran (2011)	cloze test receptive recognition test	38 L1 Arabic EFL vs. 30 ESL advanced learners	48 grammatical and lexical collocations: V+N, V+Prep, Adj+N	-learners' environment has an effect on collocation acquisition -ESL learners outperformed EFL learners - receptive knowledge was broader than productive collocations knowledge

				- collocations are a source of difficulty for the L2 learner
Fatemeh & Arshad (2011)	analysis of L2 essays c-test (a variation of cloze test)	27 L1 Persian PhD students	Lexical collocations: V+N	- the quantity, and first of all quality of collocation knowledge contributes to better writing/productive skills
Eftekhari & Rahimi (2011)	receptive recognition task reading test part of CPE (Cambridge proficiency in English)	45 Iranian EFL university students at three different proficiency levels	64 non-congruent involving common verbs (<i>have, give, take, make</i>) in delexical uses	- collocations of common verbs in delexical verbs need to be thoroughly learned, otherwise fossilization of errors occur -consciousness-raising is required for this type of collocations
Laufer & Waldman (2011)	L2 learner corpus NS corpus	L1 Hebrew and Arabic college and university students in three different proficiency levels	Lexical collocations V+N	- learners at all levels underproduced collocations in comparison to NSs -only advanced learned showed a significant increase of collocations - errors (L-induced) were persistent in similar proportion across all levels, but increases at advanced levels - collocations acquisition is a long and uneven process

*HKUST (Hong Kong University of Science and Technology) learner corpus

Appendix II: The test instrument: the collocation test with a questionnaire

A translated version in English

This questionnaire is completely anonymous and collected for strictly scientific purposes. All personal information revealed will be kept confidential and will not have any effect on your academic performance. To insure high quality of the study it is necessary that you address each item and provide only honest and truthful answers.

Thank you for your participation!

First, please answer some personal information questions about your English learning experience. In the questionnaire below check the appropriate box or provide additional information when required.

1. My mother tongue:

Ukrainian Russian other _____

2. I have been learning English since

kindergarten from _____ grade other _____

3. In school, I was in a group with intensive instruction in English

No Yes

4. Before entering university, I stayed in an English-speaking country

No

Yes less than 2 week

2-4 weeks

over a month specify _____

5. Before entering university, I was involved in English learning outside the classroom (e.g. with a parent or tutor, individually or in a private English language school). If yes, specify how intensively.

No

Yes

from _____ until _____ grade if in a private school, specify how long _____

1-2 sessions a week

60 minutes per session

more than 2 sessions a week

longer than 60 minutes per session

6. After entering university, I stayed in an English-speaking country

No

Yes less than 2 week

2-4 weeks

over a month specify _____

7. After entering university, I have been learning English outside the classroom (e.g. with a parent, tutor, or in a private English language school etc.). If yes, specify how intensively.

No

Yes

from _____ until _____ grade if in a private school, specify how long _____

1-2 sessions a week

60 minutes per session

more than 2 sessions a week

longer than 60 minutes per session

8. I am involved in learning and practicing English individually in addition to university studies in that I...

8.1**very** often read in English for improving my language knowledge, or just for pleasure

No

Not really

Yes

Specify (e.g. fiction literature, newspapers, Internet articles and news etc.)

8.2**very** often listen to English (radio/songs), or watch something (e.g. TV, movies) in English

No

Not really

Yes

Specify

8.3**very** often communicate in English with foreigners or native speakers of English

No

Not really

Yes

Vocabulary Test

The items in square brackets [] are word-for-word English translations of the Ukrainian items.

The test consists of two parts. Each part is explained in turn. You have **25 minutes** to complete the entire test. Don't spend **too** much time answering **one** item. If you are not sure about the answer, just move on to the next item. In your answers, use words **only** of general or neutral style.

Part 1 contains 20 sentences with blanks that you are to fill in with only ONE suitable word, which is a verb in the infinitive form. For most of the items, the initial letter of the missing word is provided as well as the Ukrainian translations of the target phrase in brackets at the end of every sentence. Here is an example for you!

Example:

1. I've failed and I need to t_____ the exam a second time now. [put together an exam]

Answer 1. I've failed and I need to take the exam a second time now. [put together an exam]

Test Part 1

1. Well, of course you can m_____ a profit from services! [receive a profit]
2. Then we'll p_____ a visit to your ill sister. [do a visit]
3. We will invite musicians to r_____ money for Save The Children fond. [gather money]
4. Firemen want people to wear clothes which won't c_____ fire. [become in fire]
5. If you are not satisfied, just m_____ a complaint. [give a complaint]
6. Such differences can g_____ rise to some problems in a relationship. [to cause]
7. Police finally need to b_____ the cycle of crime. [tear the vicious circle]
8. Let me pay for it. I don't want people to g_____ the impression I am poor. [impression is put together]
9. It's too soon to _____ any conclusions. Let's wait for the results. [do conclusions]
10. The oldest woman to _____ birth was 70 years old. [to bear]
11. I have been privileged to h_____ office for 11 years in a series of jobs. [occupy a position]
12. " Don't m_____ threats, " he warned her. " I don't like threats. [to threaten]
13. In any case, it will _____ no harm if you ask the teacher for some feedback. [inflict harm]
14. From a very early age we use various forms of play to m_____ sense of the world. [to comprehend]
15. The company will have to c_____ prices below \$500 to sell more. [lower the prices]
16. The program aims to r_____ awareness of global warming causes. [inform]

publicity]

17. Now Allan realized that he must _____ the burden of her death forever. [wear burden]
18. Thousands of people had gathered to p_____ their respects. [give last respects]
19. I recommend to _____ the courses at the local college. [go through a course]
20. This allows citizens to m_____ a contribution to the political process. [bring in contribution]

Congratulations! You have reached Level 2!

In **Part 2**, translate 20 phrases from Ukrainian into English. Supply only the infinitive form of the verb phrases and use **each** gap specially provided for one word/article without particle *to*. Please take a look at the example for you to help!

Example:

складати іспит: _____ [put together an exam]

Answer: складати іспит: take___ an__ exam__ [put together an exam]

Test Part 2

English target collocations	Gaps to fill in			Word-for-word Ukrainian translations
1.give a lecture				to read/lead a lecture
2.make a decision				to accept/receive a decision
3.catch a cold				one word="to become (cold) sick"
4.take a risk				to risk
5.do experiments				to experiment/ to lead experiments
6.draw the curtains				to close the curtains
7.get an opportunity				to receive a possibility
8.break a promise				to violate a promise
9.give pleasure				to bring satisfaction
10.carry a gun				to wear/bear a gun
11.give an example				to present/introduce an example
12.hold a conference				to lead/conduct a conference
13.take action				to assume measures
14.raise a question				to lift/pick up a question
15.make excuses				one word="to justify yourself"
16.pay attention				to turn/ direct attention
17.pay compliments				to say complements
18.cut production				to lower/decrease production
19.raise voice				to lift voice
20.make a mistake				to do a mistake

The original version of the test in Ukrainian

Цей тест є абсолютно анонімним і проводиться для виключно наукових цілей. Вся особиста інформація зазначена вами буде зберігатися конфіденційною і не матиме жодного впливу на вашу академічну успішність. Для забезпечення надійних результатів дослідження, очікується, що відповіді будуть максимально правдивими і відвертими.

Дякуємо за Вашу участь!!

Спочатку заповніть деяку інформацію про себе і ваш досвід вивчення англійської мови. Відмітьте галочкою відповідний квадратик або коротко впишіть необхідну інформацію у спеціально для цього відведеному місці.

1. Моя рідна мова:

Українська Російська інша _____

2. Я вивчаю англійську з

дитсадку _____ класу інше _____

3. У школі я була/був у класі з поглибленим вивченням англійської мови

Ні Так

4. До вступу до ВНЗ я перебува(в)ла в англomовній країні

Ні

Так менш ніж 2 тижні

2-4 тижні

більше місяця вкажіть скільки _____

5. До вступу до ВНЗ я займа(вся)лась англійською мовою окрім школи (з батьками, репетитором, самостійно, на мовних курсах і. т. п.). Укажіть, як інтенсивно.

Ні

Так з _____ до _____ клас якщо курси, укажіть як довго _____

1-2 рази/тиждень

60 хвилин/заняття

2 рази і більше/тиждень

більше 60 хвилин заняття

6. Після вступу до ВНЗ я перебува(в)ла в англomовній країні

Ні

Так менш ніж 2 тижні

2-4 тижні

більше місяця уточніть скільки _____

7. Після вступу до ВНЗ я займаюсь/займа(вся)лась англійською мовою позакласово/додатково (з батьками, репетитором, на мовних курсах і. т. п.). Вкажіть, як інтенсивно.

Ні

Так

з _____ по _____ семестр якщо курси, як довго _____

1-2 рази/тиждень

60 хвилин/заняття

2 рази і більше/тиждень

більше 60 хвилин/заняття

8. Додатково до мого навчання у ВНЗ я активно займаюсь англійською мовою, а саме...

8.1 ...**досить** часто і багато читаю англійської мовою просто для себе чи задоволення

Ні

Не дуже

Так

Уточніть, що саме (наприклад, книжки, газети, Інтернет і.т.п.)

8.2... **досить** часто слухаю (радіо/пісні) та /або дивлюсь щось (фільми/ТБ) англійською мовою просто для себе чи задоволення

Ні

Не дуже

Так

Уточніть, що саме

8.3 ...**досить** часто спілкуюсь англійською мовою з іноземцями і/або носіями мови

Ні

Не дуже

Так

Тест на лексику

Тест складається з двох частин. Кожна частина пояснюється по черзі. У вас є **25 хвилин** на весь тест. Не затримуйтеся дуже довго на **одному** пункті. Якщо точно не знаєте або не пам'ятаєте відповіді, ідіть далі. Не забувайте, цей тест не на оцінку. У відповідях використовуйте *лише* лексику загального/нейтрального стилю (будь ласка ніякого сленгу!).

У першій частині є 20 речень з пробілами куди потрібно вписати дієслово, яке найбільше підходить за смыслом в інфінітивній формі. У якості допомоги дана початкова літера відсутніх дієслів (не у всіх реченнях!), а також переклад фрази у дужках вкінці речення. Дивіться зразок унизу.

Зразок

1. I've failed and I need to t_____ the exam a second time now. (складати іспит)

Відповідь: 1. I've failed and I need to take the exam a second time now. (складати іспит)

Частина 1

1. Well, of course you can m_____ a profit from services. (отримувати прибуток)
2. Then we'll p_____ a visit to your ill sister. (робити візит)
3. We will invite musicians to r_____ money for Save The Children fond. (зібрати кошти)
4. Firemen want people to wear clothes which won't c_____ fire. (загорятися)
5. If you are not satisfied, just m_____ a complaint. (подати скаргу)
6. Such differences can g_____ rise to some problems in a relationship. (викликати)
7. Police finally need to b_____ the cycle of crime. (розірвати порочне коло)
8. Let me pay for it. I don't want people to g_____ the impression I am poor. (складається враження)
9. It's too soon to _____ any conclusions. Let's wait for the results. (робити висновки)
10. The oldest woman to _____ birth was 70 years old. (народжувати)
11. I have been privileged to h_____ office for 11 years in a series of jobs. (займати посаду)
12. "Don't m_____ threats, "he warned her." I don't like threats. (погрожувати)
13. In any case, it will _____ no harm if you ask the teacher for some feedback. (завдавати шкоди)
14. From a very early age we use various forms of play to m_____ sense of the world. (розуміти, осягати)
15. The company will have to c_____ prices below \$500 to sell more. (знизити ціни)

16. The programme aims to r_____ awareness of global warming causes.
(проінформувати громадськість)
17. Now Allan realized that he must _____ the burden of her death forever. (нести тягар)
18. Thousands of people had gathered to p_____ their respects. (віддати останню шану)
19. I recommend to _____ the courses at the local college. (пройти курс)
20. This allows citizens to m_____ a contribution to the political process. (вносити вклад)

Вітаємо! Ви закінчили перший рівень!

У **другій** частині тесту потрібно перекласти 20 фраз з української мови на англійську. Необхідно написати вирази в інфінітивній формі **без** частки **to**, але заповнити кожний відведений для цього пробіл. Дивіться зразок унизу.

Зразок

складати іспит: _____

Відповідь: складати іспит: take an exam

Частина 2

Українські вирази	English translation		
читати/проводити лекцію			
приймати рішення			
застудитися			
ризикувати			
проводити експеримент			
закрити вікно занавісками			
отримати можливість			
порушувати обіцянку			
приносити задоволення			
носити зброю/пістолет			
наводити приклад			
проводити конференцію			
ужити заходів			
піднімати питання			
виправдовуватися			
звертати увагу			
робити комплементи			
скоротити виробництво			
підвищувати (чийсь) голос			
помилятися/робити помилку			

Abstract in German

Die vorliegende Arbeit beschäftigt sich mit der Entstehung von Verb-Substantiv-Kollokationen bei ukrainischen/russischen Muttersprachlern die Englisch als Fremdsprache lernen und sich bereits auf einem fortgeschrittenen Niveau befinden. Die Bedeutung von Kollokationen im Spracherwerb und Sprachunterricht findet eine weitverbreitete Anerkennung. Es liegen wissenschaftliche Beweise vor, dass solide Kollokationskenntnisse der Zielsprache die Leistungsfähigkeit in L2 bedeutsam erhöhen, sowie zu fließendem und idiomatischem Sprachgebrauch beitragen (Bahns & Eldaw 1993; Howarth 1996; Pawley & Syder 1983). Außerdem unterstützen Kollokationen den allgemeinen Spracherwerb, in dem sie sowohl rezeptive als auch produktive Fertigkeiten fördern (Bardovi-Harlig 2002; Lewis 1993; Nattinger & DeCarrico 1992; Wray 2000).

Trotz der vielfachen Vorteile dieses Aspektes des Wissensstandes, hat die Forschung immer wieder gezeigt, dass L2 Kollokationen eine ernste Herausforderung an Nichtmuttersprachler darstellen (e.g. Bahns & Eldaw 1993; Biskup 1992; Bonk 2001; Gitsaki 1999; Howarth 1996; Laufer & Waldman 2011; Nesselhauf 2005). Ein Sprachschüler tendiert zu einem Mangel an L2 Idiomatik und einer Fülle an Kollokationsfehlern, die bis zu fortgeschrittenen Sprachniveaus vorherrschen und nicht den Anschein erwecken, sich trotz extensiven L2 Inputs zu ändern. Obwohl mehrere Studien zu L2 Kollokationskenntnissen und deren Gebrauch durchgeführt wurden, ist die Forschung, die sich darauf konzentriert, wie Kollokationen in einem fremdsprachigen Unterricht, insbesondere von ukrainisch/russisch sprechenden Schülern erworben worden sind, sehr rar. Deshalb kann die vorliegende Studie als ein wertvoller Beitrag zu diesem Forschungsfeld betrachtet werden, indem sie versucht, die vorhandene Lücke zu schließen und Kenntnisse über den L2 Kollokationserwerb im Allgemeinen zu erweitern.

Genauer gesagt hatte die Studie zwei wichtige, miteinander verbundene Ziele. Das erste Ziel war, die produktiven Kollokationskenntnisse von Studenten zu bewerten, und ihre Entwicklung vom ersten bis zum letzten Studienjahr zu verfolgen. Das zweite Ziel war, Faktoren zu erforschen, die das Kollokationslernen beeinflussen könnten. Hierbei wurde besonderer Wert auf die Frequenz von Kollokationen, die Quantität (Länge) und die Qualität (Intensität und Typ) von EFL-Inputs gelegt. Um die Wirkung des Typs der Qualität, d. h. den Einfluss des unmittelbaren EFL-Unterrichts zu erforschen, wurde Kollokationsleistung zwischen Studenten von zwei verschiedenen EFL-Unterrichtsumfeldern (Klassentypen) verglichen. Ein sekundäres Ziel war auch, Vorschläge für Kollokationsunterricht zu formulieren, basierend auf den erhaltenen Ergebnissen.

Die Analyse spezialisierte sich auf produktive Kenntnisse von inkongruenten (nichtäquivalente L1-L2) Verb-Substantiv-Kollokationen, die allgemeine Verben im nichtlexikalen Gebrauch einschließen. Dieser Kollokationstyp wird in der Literatur als besonders schwierig für L2 Schüler betrachtet, und verursacht gleichzeitig eine beträchtliche L1-Übertragung, die zur Produktion von abweichenden Kombinationen führt. (e.g. Biskup 2001; Howarth 1996; Laufer & Waldman 2011; Nesselhauf 2005; Yamashita & Jian 2010). Die Teilnehmer der Studie waren insgesamt 80 fortgeschrittene L2 Studenten von zwei verschiedenen EFL-Unterrichtsumfeldern. Sowohl die Gruppe im ersten Jahrgang als auch die im vierten bestand aus 20 Studenten. Die betreffenden EFL-Unterrichtsumfelder sind zwei unterschiedliche Studienrichtungen: Übersetzung/Dolmetschen und Englisch Lehrer-Ausbildung, die sich in zwei verschiedenen Institutionen zweier ukrainischer Städte, Tscherkassy und Kiew befinden. Um die Ziele der Studie zu erreichen, wurde ein spezielles Testinstrument entwickelt. Der Test bestand aus einem 20-teiligen Cloze-Test und aus einer 20-teiligen Übersetzungsaufgabe. Um den Einfluss der Frequenz von Kollokationen beim Erwerb zu erforschen, waren die in den Test eingeschlossenen Verb-Substantiv-Kollokationen von zwei verschiedenen Frequenzbändern: Hochfrequenz- und Niederfrequenz.

Ergebnisse der Analyse weisen darauf hin, dass nichtkongruente nichtlexikale Verb-Substantiv-Kollokationen eine ernste Herausforderung an ukrainisch/russisch sprachige EFL Studenten darstellen, sogar für Studenten eines fortgeschrittenen Niveaus. Obwohl der Erwerb von Verb-Substantiv-Kollokationen im EFL-Unterricht implizit erfolgt, findet die Aneignung ziemlich langsam und unvorhersehbar statt. Insbesondere die beobachtete mangelhafte Kenntnisse von Kollokationen der Teilnehmer und das geringe Wachstum von Kollokationen vom ersten bis zum letzten Studienjahr, können herangezogen werden, um das gesamte niedrige Bewusstsein von EFL Studenten für L2 Kollokationen als legitime und wichtige lexikalische Elemente hervorzuheben. Faktoren die gefunden wurden den Erwerb der Zielkollokationen positiv beeinflussen sind: die Frequenz der Elemente, die Qualität des L2-Inputs und außeruniversitäres Lernen der EFL Schüler, mit der vorrangigen Wichtigkeit des letzteren Faktors. Alles in allem zeigen die Ergebnisse der Studie nachweislich, dass damit besonders schwierige Kollokationen im EFL-Unterricht effektiv erlernt werden können, unbedingt ein reichhaltiger und umfassender Kontakt mit der Zielsprache (implizite Techniken) sowie spezielle pädagogische Behandlung (explizite Techniken) sind erforderlich.

Curriculum Vitae

Persönliche Daten:

Name: Maryna Dudnyk
Geburtsdatum: 09. Juli 1981
Geburtsort: Tscherkassy, Ukraine

Ausbildung:

ab 10/2008 **Universität Wien**
Bachelorstudium: Transkulturelle Kommunikation
(Russisch, Deutsch, Englisch)
Masterstudium: English Language and Linguistics

09/2001 - 07/2005 **Staatliche Technische Universität Tscherkassy**
Abschluss: BA (Englische Philologie) mit Auszeichnung

09 - 12/2004 **Volkshochschule**, Deutschland, Deutschkurs (Mittelstufe)

06 - 09/2005 **Volkshochschule**, Deutschland, Intensivdeutschkurs (Oberstufe)

09 - 12/2005 **Languages and Teacher Training Centers**, Dnepropetrovsk,
Ukraine; Moskau, Russland, Teacher Training Courses

09/1998- 07/1999 **Istanbul Universität**, Türkei, Türkischkurse
(Grundstufe -Mittelstufe)

09/1995 - 07/1998 **Tscherkassy Gymnasium**, Ukraine, Maturaabschluss

09/1987 - 07/1995 **Allgemeinbildende Musikschule**, Tscherkassy, Ukraine

Berufliche Weiterbildungen:

ab 2005 Selbstständige Tätigkeit: Übersetzen/Dolmetschen,
Nachhilfestunden (Russisch, Ukrainisch, Englisch, Deutsch)

ab 09/2007 Gründung von Non-Profit-Organisation, Ausbildungszentrum
„Evrika“:
Organisationsleitung, Unterrichten, Übersetzen/Dolmetschen

08/1998 – 01/1999 Hostess im Kulturunterhaltungszentrum „James Joyce“, Istanbul,

Türkei

01/1999 - 07/2000 Sachbearbeiterin/Dolmetscherin in der Reiseagentur „Alba Turism“,
Istanbul, Türkei

Weitere Kenntnisse:

Sprachen: Russisch (Muttersprache)
Ukrainisch (ausgezeichnete Kenntnisse)
Englisch und Deutsch (sehr gute Kenntnisse)
Türkisch (gute Kenntnisse)
Französisch (Grundkenntnisse)