Acquiring the English Collocations with some Element of Unpredictability

by means of chunking, comparing, guessing

Taťána Pokorná
Kristianstad University College
The School of Teacher Education
English IV, Spring 2009
D-essay in English Didactics
Tutor: Carita Lundmark

Contents

1 Introduction

Language knowledge is collocational knowledge. Humans can communicate without severe difficulties even when they use words which are placed in the wrong order, their pronunciation is not perfect or marked with the proper grammatical morphemes; however, communication tends to break down if we do not use the correct word in terms of meaning. Moreover, the 'correctness' of those words is further modified by other syntactical components they generally occur together with in the discourse.

For example, native speakers whose mother tongue is English would probably never “wear *rancid socks* or find *fetid milk*”(Aitchison 2003:13). Why do not the English *concentrate to* their work? Or *could not eat a cow* when they feel extremely hungry? Common word co-occurrence in language seems to cause genuine misunderstandings for second language learners, particularly when they have attained a certain level of proficiency.

The second language learners encounter a great number of different words on their long and hard way through acquiring a foreign language (L2). Those learners, as it is assumed, attempt to not only translate an L2 word into an L1 word and vice versa, but also be able to utter them in a native-like way to be at least easily understood.

If the students of an L2 have an excellent command of language, particularly in terms of collocations, they are capable of expressing every single feeling or emotion by using spoken or written language. Moreover, their language performance can hardly be interfered with ambiguous phrases, as they are able to see the difference, for example, between *light wine* and *white wine* or *cooked beer* and *brewed beer*. With regard to these facts, the question arises as to how the teachers of an L2 can facilitate learners' acquisition of such a crucial knowledge of language.

1.1 Aim

The aim of this research is to investigate how learners of English as a second language acquire one of the most crucial aspects in language, collocations. Some aspects of collocational knowledge in a second language acquisition are focused on:

1. whether the learners remember much easier if content words collocate with content words than when function words collocate with content words;
2. whether the learners acquire collocations like one word and analyse them afterwards or vice versa;
3. whether the rhythm of language can help the learners with learning L2.

This research also intends to investigate the learners' opinions, feelings and experiences that have shaped the way the learners face up to and deal with collocations in a second language. Since there are both women and men investigated more in a qualitative than quantitative way, there might appear some gender differences. The last intention is not the central one.

1.2 Material and Method

1.2.1 Participants

Learners' approach to acquiring collocations of an L2 will be investigated. The research will take place at a university in southern Sweden. Six students (four women and two men) will be individually interviewed by the researcher. It will be an informal discussion between the researcher and the participants. Every investigated participant studies in Sweden as an exchange student, which means that their mother tongue is not Swedish. The age of the participants is between 20 and 26. They have reached at least the intermediate level in English, which is the basic level for starting to concentrate on learning unpredictable collocations (*give up*, *severe traffic jam*) consciously, since these learners have sufficient vocabulary.

The level of the participants ought to be intermediate because all of them are the exchange students studying at universities in the Czech Republic, Spain, Russia and Austria. Therefore, it is assumed that they were asked by coordinators at their home universities to write a formal letter in English that had to consist of reasons for being chosen as an exchange student. Furthermore, three of the participants have passed the Cambridge exam First Certificate in English. The participants were chosen at random and none of them are familiar with the topic of the research.

To make the participants feel comfortable and unstressed is one of the reasons why they are not informed in advance about the topic of this research. They might intend to start studying, for instance, idiomatic language by heart and feeling under pressure that they would not know some answers. The next ground for not mentioning the term collocation is that the participants who have not heard of such a word before, might go and look its definition up. In that case the interview question *Have you ever heard of the term collocations?* would be irrelevant. This question should reveal whether the participants care about linguistic terminology.

Therefore, the participants have obtained the information that there are two short tests which are comprised of the interview itself. Additionally, they have received three more nuggets of information:

1. the research is focused on some ways of acquiring vocabulary items in English,
2. they will be asked to comment on a few linguistic features,
3. the interview might last about an hour or two and will be recorded on the Dictaphone.

Six different criteria in terms of students' characteristics are taken into account to create the short tests: age, nervousness, interests, social background, general knowledge, preparation (see Table 1 below). Such an idea has been inspired by Bachman and Palmer (1996) who suggest that "[p]ersonal characteristics are individual attributes that are not part of test takers' language ability but which may still influence their performance on language tests" (Bachman & Palmer 1996:64).

Table 1: Six Criteria and Participants' Characteristics

Age	From 20 to 26
Nervousness	Even though the participants are not acquainted with the reason for their testing and with the topic of this research (only before doing the short test), it is supposed that they will not feel nervous because they will at least know that there is no time limit for finishing the short tests.
Social background	They come from a large number of different countries with their specific attitudes towards life and education.
Interests	As the age of the students is specified only by the fact that they belong to the group of adults (may be teenagers), it is difficult to specify students' interests. On the other hand, they have to be able to study at university, which refers to the knowledge of academic vocabulary.
General knowledge	Those students had to finish schools which have provided them with the specific knowledge needed for studying at university. They are supposed to have achieved a level of education essential for university.
Preparation	The students are expected not to be prepared for some specific fields of vocabulary.

1.3 Interview

To find out how learners acquire collocations in English interview questions (see appendix C) will be prepared. Those questions will be divided into two sections. One section is focused on linguistic features of collocations (this part is mainly based on the tests) and another section is based on socio-cultural aspects of collocations in language.

Two short tests are prepared for the participants. The tests are focused on both receptive (multiple-choice) and productive (key-word transformation) skills. The reason is that the participants are asked which one they feel is more helpful to acquire collocational expressions and why. Both the tests help with the preparation of interview questions.

1.3.1 Tests

Each test consists of five samples of collocations and has been designed on the basis of the theoretical background. The multiple-choice test comprises one sample of idiomatic expressions that are called binomials, three examples of the common word co-occurrence in type: two content words collocate with other content words and one content word collocate with a function word. The last sentence consists of a sample of a link between concepts in the mental lexicon.

The key-word transformation test is compiled of two samples of co-occurrence in type (two content words collocate with two function words). There is also an example of the grammatically and lexically fossilized word co-occurrence which is followed by a content word collocating with a function word. The last two samples represent two types of idiomatic expressions based on metaphorical language.

To find some examples of frequent word co-occurrence the Corpus of Contemporary American English (Davies 2004) was used. In the tests there is a sample of collocations that is common particularly in fiction, however it does not frequently occur in spoken English. The tests also comprise the examples of collocations which are highly frequent in newspapers, magazines and academic fields.

1.4 Procedure

Having answered the first four questions of the interview (which focus on the participants' feelings and opinions on English collocational knowledge), every participant is asked to concentrate on the samples of collocations in two short tests: the multiple-choice test which tests receptive skills, the key-word transformation test which tests productive skills. A learner is interviewed by the researcher while he or she does the tests. Most of the interview questions are prepared in advance on the basis of the theoretical background. Every single interview is recorded on the Dictaphone, since it is undesirable to miss something that each student utters.

1.4.1 Interviewing

Stages:

1. The interview starts with some general questions that focus on the participants' knowledge of English collocations.

2. The multiple-choice test is presented orally at first because it might help the participants find the correct answer in case that a participant feels insecure about the written form. The participants are asked to listen to the researcher who will be reading five sentences with blanks that should be filled in by the listeners. After every single sentence, the five alternatives are given to the participants. Furthermore, this sentence is shown to an interviewee in writing to enable him or her to think about the sentence more deeply, and then reply to the researcher's questions. The main task of the researcher is to jot down the participants' answers with no correction.

3. Having done the multiple-choice test, the participants do the key-word transformation test. Similarly to the multiple-choice test, they have so much time they need to finish the test. If they do not know the meaning of a word, it is allowed to look it up in the Google Dictionary on the Internet. The reason is that the aim of this research is not to investigate the capacity of the mental lexicon. The participants are interviewed afterwards. Some interview questions might be modified as the interviews proceed. This fact means that there will be six quite different linguistic parts of the interview .

4. The interviews continue to the second section. The participants reply to the questions focused on their personal feelings, experiences and opinions. They are expected to talk about how different is their current learning style from the one they applied when they started learning English, for instance.

5. Finally, when the researcher has collected all data on the Dictaphone, every single interview will be analysed in terms of linguistics and pragmatics.

2 Theoretical Background

“Most everyday words do not have an independent meaning, or meanings, but are components of a rich repertoire of multi-word patterns that make up text.“(Sinclair 1991:108) In other words, if we use the word *ball* separately with no other words around, our listeners might probably imagine a small/big round object at first. However, subsequently their imagination would go further to other meanings and the communication would break down.

On the other hand, such utterances as *I passed the ball yesterday* and *I don't want to go to the ball* convey the different meanings of these homonyms as explicitly as possible. The reason is that *pass the ball* is a collocation which refers to *throw* or *kick the ball to someone else.* Consequently, *throw the ball to someone* has quite different meaning from *throw the ball at someone.* Since the meaning of every single word is modified by its occurrence in the specific context, particularly by its co-occurrence with other vocabulary items, the teachers' duty is to enable learners of an L2 to encounter words in as many different contexts as the school lessons allow.

2.1 Collocation

2.1.1 Content vs Function Words

Content words carry the meaning which is open for newcomers. On the other hand, the main task of function words is to organize those meanings. According to Harley, content words “carry the meat of the message and function words are responsible for providing the structure that let us communicate better than Tarzan” (Harley 2006:186). Metaphorically, Aitchison (2003:152) likens the relationship between content and function words to the connection between bricks and mortar. To be specific, bricks are meaningful if they are fixed together with mortar until the building stands.

As for the second language acquisition which takes place in a classroom, it is the teachers' duty to enable students to encounter an L2 as one elaborate system where everything is interconnected, since this is the principle of the mental lexicon. People can be understood

even when they utter “so called telegraphic sentences” (Lightbown & Spada 1999:2) but it is really exhausting for listeners to be a part of such a difficult conversation. No one might prefer to be seen as Tarzan in a society.

For this study it seems to be useful to divide content and function listemes according to word classes. Content words refer to the nouns (*theatre, spinster, love, water*), verbs (*spring, fall, seem, adjust*) and adjectives (*considerable, tall, red*), which are flexible words either in form or meaning. On the other hand, the meanings of the function listemes are pretty inflexible. To define flexibility and inflexibility of the meaning of every single word, Harley's study suggests that “when a function word and content word don't go together naturally, the one whose meaning changes to accommodate the meaning of the other is the content word, not the function word” (Harley 2006:212). In other words, phrasal verbs can represent such a case, since the content word *set* is forced to accommodate its meaning to the function word *about* when the phrasal verb *set about (sth)* is taken into account. The meaning of *set* is dependent on *about*. Some examples of English function words can be conjunctions (*but, and*), prepositions (*at, in, to*) and definite and indefinite articles (*a, the*).

2.1.2 How is the Term Collocation Defined?

Oxford Collocations Dictionary (2002) deals with a broad but from the dictionary point of view satisfactory definition: “Collocation is the way words combine in a language to produce natural-sounding speech and writing” (OCD 2002). Two word combinations are used to exemplify this phenomenon, namely *strong wind* and *heavy rain*. *Heavy wind* and *strong rain* would be taken as awkward combinations by native speakers. Nation (2001:56) treats those samples as co-occurrence in type. To be more specific, the content words collocate with the content words.

According to Sinclair (1991), there are two different principles of interpretation of word occurrence in a text. The first one, “the open-choice principle”, is a way of seeing language text as the result of a very large number of complex choices […] the only restraint is grammaticalness” (Sinclair 1991:109). If this statement were true, then the word combinations such as *rancid socks* and *fetid milk* or *heavy wind* and *strong rain* would never be felt to be incorrect by a native speaker, since the grammatical pattern that an adjective can modify a

noun would be followed. Specifically, in such cases as *heavy wind* and *strong rain*, both the adjectives occur in "attributive function" (Quirk *et. al* 1990:129), which is one of the main criteria of using adjectives according to grammatically correct patterns.

The open-choice principle is based on Chomsky's theory. According to Chomsky (2006), a person does not have "a stock of utterances that he produces by 'habit' on an appropriate occasion […] [n]or […] [does] the speaker [have] a stock of 'patterns' in which he inserts words or morphemes" (Chomsky 2006:106). Such an explanation is related to what Chomsky (2006) calls a linguistic 'competence', which is the system of rules in terms of grammar. Human utterances would only be restricted by grammatical rules, sentences would be infinite if people did not have so called extralinguistic beliefs that Chomsky (2006) labels as 'performance'.

Figuratively, there is no specific pattern that would determine the concrete length of sentences or how often people should use some specific expressions. On the one hand, with regard to competence, it is possible to connect *heavy* with *wind* because a grammatical rule is kept. On the other hand, when performance is taken into account, *heavy* cannot be linked with *wind* since, as Chomsky (2006:102) emphasizes, people do not interpret the actual utterances simply by applying the linguistic principles.

Apparently utterances differ from person to person. However, they are composed by means that are "the socioculturally meaningful artifacts and symbolic systems of a society, the most important of which is language" (Zuengler & Miller 2006:39). It follows that it is possible to use *write a novel* in different contexts with the meaning that someone uses a pen or pencil (currently the keyboard of a PC) and produces letters or numbers on a surface. It is treated as a norm of a society. *Invent a novel* might be seen, however, more problematic. It may refer to either that someone writes this kind of genre for the first time or metaphorically that someone writes a novel by inventing stories and plots.

From the point of view of teaching a second language, the open choice principle cannot be accepted since language belongs to other symbolic systems, such as traffic signs, by which language is influenced. It is a social construct. Moreover, second language learners modify their utterances according to the amount of exposure. They particularly use language in order to interact and communicate meanings, as Ellis (1997:122) emphasizes.

The other principle of interpretation of word occurrence in a text, the so called 'idiom principle' (Sinclair 1991:110), is based on the fact that:

> [W]ords do not occur at random in a text […] a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments. (Sinclair 1991:110)

For instance, the word combination *of course* is probably stored in the mental lexicon as one single word. When we analyse this phrase 'word by word', then we find out that *of* generally operates as a preposition or quantifier with nouns on its both sides. On the other hand *course* belong to the content words, which means that it should carry the meaning of the phrase *of course.* In fact, the use and meaning of *course* and *of course* are completely different.

Nation's study (2001) suggests that "it is not sufficient to define a collocation as a group of words that frequently occur together" (Nation 2001:324), since there are some groups of words in terms of corpora which appear frequently together (*although he*, *but if*). However, those words do not change their meaning when they are used separately. From the second language learners' point of view, it would be extremely hard to learn every most frequent co-occurrence in a language by heart. Therefore, Nation (2001) defines collocations as "closely structured groups whose parts frequently or uniquely occur together […] they contain some element of grammatical or lexical unpredictability or inflexibility" (Nation 2001:324). Such lexical unpredictability, for instance, can be caused by the fact that second language learners are accustomed to connecting some words differently in their mother tongues than in a second language (in this study English).

Nation's (2001) definition is highly relevant for this study. On the basis of this definition both the tests are designed. Particularly the unpredictability and inflexibility of the collocations in the tests are taken into account. In addition, they belong to the high frequency of occurrence. For example, it is possible to say or write *heavy traffic* and *severe traffic jam* but not *big traffic* or *big traffic jam*, which are very frequent errors of learners of English from the Czech Republic.

2.1.3 The Division of Collocations

It might be helpful to mention Nation's (2001:56) division of collocations according to the size and type in which they occur in the discourse or how many words in terms of meaning can modify one 'node', which is described by Müller (2008:4) as a word in the middle of other words. These words are somehow connected to such a word in the middle and are called 'collocates', like in a mind-map. For example, when *coffee* is labelled as a node, the words *strong, bitter, drink* are its collocates. As far as type is concerned, there are two different types of listemes, content and function. In relation to collocations, function words can collocate with content words (*wait* with *for*) or content listemes can collocate with content ones (*devastating* with *flood*).

From the point of view of 'the possible range of collocates and size' (Nation 2001:56), the learners of an L2 tackle the difference between metaphorical (figurative) language and common word co-occurrence. For example, idioms are often created by means of figurative language (*You seem to be on cloud nine today.*) while academic discourse presumes some certain knowledge of common word co-occurrence (*commit a crime, suicide* but not *an accident*). The term collocation is comprised of both these features of lexicology.

2.1.3.1 Two types of frequency

The frequency of co-occurrence should be also taken into account when collocations are discussed. Kjellmer (1984), as mentioned in Nation (2001:329), deals with two different terms such as absolute frequency and relative frequency. Absolute frequency refers to “the actual number of times a collocation occurs in a corpus. Relative frequency compares actual frequency of occurrence with and expected number of occurrence” (Nation 2001:329).

For instance, the collocation *heavy traffic* can be treated as one whole expression and then we search for the absolute frequency. However, when this collocation is analysed item by item, we find out that *traffic* can be connected to some other collocates as well as *heavy*. Namely, OCD (2001) deals with a great number of adjectives (*congested, constant, busy* etc.) which

can collocate with the noun *traffic*. In this case, we talk about the relative frequency. More details are discussed in the third section of this research.

2.1.4 Thinking in Concepts

Our mental lexicon is conceptual in nature. People think in concepts that "are organized bundles of stored knowledge which represent an articulation of events, entities, situations, and so on in our experience" (Cruse 2004:125). Such bundles are interconnected like a complex web by links which can be of different kinds (*the, a kind of, a shred of, is* etc.). The connection between them can be either direct or indirect. How directly or indirectly those bundles are linked depends on people's everyday experience. For example, according to Müller (2008:4), the word *pig* can be easily associated with the word *pork*. Yet, a farmer who breeds pigs may associate such a word with a much larger number of concepts (*feed, domestic, piglet, boar, sow* etc.).

As far as collocations in second language learning are concerned, when learners from the Czech Republic encounter the 'binomial' (McCarthy & O'Dell 2001:164) *odds and ends* for the first time, they will not probably have to create a new concept, since they can have created one some time before by means of their mother tongue. Those learners 'only' take and compare the meaning of English *odds and ends* with Czech *zbytky.* However, even though the learners are successful in remembering a new name of the concept, they still might not know what and how it is linked with other concepts in the English language.

It is obvious that the Czech and English concepts are linked differently. Moreover, the Czech concepts are not the same as the English ones. It follows that a second language learner must sooner or later encounter an English concept (e.g. *stile*) that does not exist in the Czech language. In that case, a completely new concept with its specific meaning has to be created in the mental lexicon. It takes much time to make connections between a new concept and well established concepts.

As far as collocations are concerned, a second language learner is probably forced to reorganize somehow his or her mental lexicon on the basis of the collocational system of an L2 (English in this study).

2.2 Metaphorical Language

Our everyday language consists of figurative language. One major part of such a language concerns metaphors. Metaphorical language is particularly based on the fact that the meanings of words, content words primarily, may differ according to how they are used in context. As Jackson and Amvela (2000:59) remarks:

> [A] word may have both a 'literal' meaning and one or more 'transferred' meanings […] metaphor, which is the most familiar kind of transference, before turning to other kinds of transference. The basic difference between metaphor on the one hand and the other types of transference on the other is that metaphor is 'irregular', because it applies to individual lexical items, whereas the other kinds may be considered more 'regular', in the sense that they do not apply just to individual lexical items but to several members of a specific class, e.g. a group of nouns or adjectives. (Jackson & Amvela. 2000:59)

Similarly, Aitchison (2003:41) discusses the difference between fixed and fuzzy meaning of words. Compared to Jackson and Amvela (2000), Aitchison's fixed meaning refers to literal and fuzzy means transferred. An example of transferred meaning and also metaphor might be the utterance *You are a big child* if the second part *a big child* refers to an adult who behaves like a child. It is a metaphor because it is irregular. There are no entries of this word co-occurrence in OALD (2005) which would offer the adequate definition of the sentence *You are a big child*. Its meaning is dependent on at least socio-cultural context and the speaker's personal experience. On the other hand, this utterance can also carry the literal meaning because the speaker may mean that a child has grown up and is older than he or she was, as OALD (2005) depicts.

Since metaphors are irregular, they should be left out of the field collocations. Yet there is a great number of word co-occurrences that are based just on a transferred meaning (*weak/strong coffee, red hair, white wine, heavy traffic, light beer*). These samples, when we use White's terminology (1996:17), are composed of 'a primary vocabulary' that is literal and non-metaphorical, and of 'a secondary vocabulary' that introduces the metaphorical language. Despite Jackson and Amvela (2000) treat the word combination *white wine* as regular and therefore as belonging to the other types of transference, this study does not separate such expressions from the term metaphor. The reason is that they are still based on metaphorical language. Therefore, a more regular metaphor cannot be separated from collocations.

According to Aitchison (2003:149), metaphors are created because either our mental lexicon is metaphorical in nature or people are influenced by socio-cultural background. With regard to how a norm influences people's metaphorical thinking, there are some metaphors taken as nice and more satisfactory than others, they are prototypical. Such figures might become clichés (*a trouble shared is a trouble halved*), idioms (*kick the bucket*) or fixed phrases because of their high frequency of occurrence in everyday language.

As far as a second language learner is concerned, understanding and producing metaphorical language might seem like an insurmountable boundary. The reason for such a feeling probably springs from the fact that metaphors are social constructs. They are created by means of the current world in which people live. Therefore, a second language learner of English from Czech who, for the first time, encounters the 'simile' (McCarthy & O'Dell 2001:162) *as cool as a cucumber* must misunderstand its meaning since the Czechs use quite a different simile, namely *as cool as a dog's muzzle.*

2.2.1 Idioms

While Nation (2001) treats idiomatic language as a part of collocations, Jackson and Amvela (2000:65) consider an idiom as a type of multiword lexeme due to its unpredictable meaning from its components. This study follows Nation's classification (2001: 329-32) of collocations. According to this classification, collocations can be divided into ten different categories and most of them comprise idiomatic language. For example, there are idioms that are grammatically (*part and parcel*) or lexically fossilized (*a bird's eye view*), which means that they do not allow any changes in terms of word order or replacing words by other words. Some idioms belong to the category called semantic opaqueness (*of course*). The meaning of such an idiom is not deducible from its parts.

There are also some idioms which carry either the literal meaning or the transferred meaning. For example, the sentence *She is always talking behind my back* may be interpreted in two ways. On the one hand, there could exist a woman who is always talking while I am standing with my back to her. On the other hand, there might be a woman who is always saying some negative things about me when I am not there. The first thing which a hearer/reader has to do is to identify whether the meaning of such a phrase is literal or idiomatic.

Jackson and Amvela (2000:66) deal with three features that are typical for idiomatic language. The first one focuses on the fact that as soon as a hearer/reader treats a part of speech (which can carry either idiomatic or literal meaning) as idiomatic, he or she automatically eliminates the literal meaning and attempts to understand an idiomatic meaning. The second essential feature of idioms concerns the importance of the context in which the idiom appears since such context helps with the reduction of potential ambiguity. The third and last crucial feature is based on the fact that idioms have special syntactical patterns. To be more specific, we cannot replace *the bucket* with *the table* in the idiom *kick the bucket* without changing the meaning. In those cases, such expressions as *kick the bucket* or *someone is always talking behind someone's back* behave like one-word expressions. They cannot be separated.

Finally, it is important for this study to explain the terms binomials and similes. Such terms are generally used in some English textbooks focused on vocabulary learning. McCarthy and O'Dell (2001:164) describe binomials as two words that are joined with a conjunction and whose order is unchangeable. For example, it is impossible to change the order of the expression *head or tail,* we cannot say *tail or head.* As far as similes are concerned, these expressions are based on metaphorical language, for instance *as thin as a rake*. Similar to binomials some of them are fixed. Yet, in that case, as McCarthy and O'Dell (2001:162) remark, they are usually treated as informal and humorous, thus a second language learner ought to use them with care and attempt to acquire them as one word with its specific meaning.

2.3 Chunking

Nation (2001:319) deals with Miller's study (1956) which distinguishes bits of information from chunks of information. Ellis (2001) proposes that “a chunk is a unit of memory organisation, formed by bringing together a set of already formed chunks in memory and welding them together into a larger unite”(Ellis 2001:39). Humans' short-term memory is limited to the ability of remembering around seven bits of information. To overcome this limitation people form those bits into meaningful chunks to be able to store them in the long-term memory. For example, the idiom *flying saucer* can be divided into two separate words

flying and *saucer*. Both the words can be further parsed into morphemes and then into sounds (in speaking) or letters (in writing).

Although there are not identical utterances, since the mental lexicon develops on the basis of human experiences, the capacity of the human mental lexicon is not infinite. There is a finite number of English phonemes and in spite of the fact that we cannot exactly say how many words or morphemes English has, it is apparent that words are fewer than simple sentences etc. However, as Nation (2001) proposes, "as chunks become bigger, their frequency of use becomes lower" (Nation 2001:321). Furthermore, if the frequency of expanded collocations is so low, they had better not be acquired as chunks.

Chunking moves in two directions: from the smallest part of speech to the largest part of speech or vice versa, which means that second language learners might acquire the phrase *Nice to meet you,* at first as one unite and later, when they reach a certain level, this phrase can be analysed. However, some learners may prefer the direction where "smaller chunks can be grouped into larger chunks" (Nation 2001:319). In other words, such learners probably acquire the phrase *Nice to meet you* word by word in terms of meaning. Namely, they may search for the meaning of *nice* then *to meet* and finally *you.* Both of these directions have their pros and cons.

Considering second language learners, one obvious advantage of chunking is that these learners are able to acquire some frequently used phrases, such as *What's your name?, What's that?, Nice to meet you., Glad to see you.* etc., as one inseparable unite. This fact enables second language learners to store and produce native-like language. Ellis (1997:129) gives some examples of grammatically correct expressions but in terms of 'native-like control of the language' they sound odd and foreign:

1. I wish to be wedded to you.
2. Your marrying me is desired by me.
3. My becoming your spouse is what I want.

According to the open-choice principle, explained by Sinclair (1991:109), such sentences are grammatically correct and thus they should be seen as 'normal'. Actually, *I want to marry you* is more appropriate and idiomatic. Compared to a first language learner, it is the similar process in acquiring language since, according to Lightbown & Spada (1992:5), children

discover language in their early years by repeating the most frequent words, or phrases, they encounter in everyday life.

On the other hand, there are some considerable disadvantages of chunking. The major drawback is the limited capacity of our memory. It might be impossible to remember language by means of chunks, since there are much more chunks than their components. It follows that "[if] chunks are learned as unanalyzed units […] the parts of unit are not available for creative combination with others" (Nation 2001:320). Concerning collocations, even though a second language learner is able to use the expression *never count your chickens until they're hatched,* he or she might not be able to use the word *hatch* in a different context.

2.4 Discourse

Everything that people say, write or do in terms of language can be classified as a realm of discourse. Dijk (1997:2) divides discourse into three dimensions: language use, the communication of beliefs (which belongs to cognitive processes) and interaction in society in which people appear or live. Additionally, these three dimensions must be further divided into two different modes, speech and text, since there is a significant difference between spoken and written forms of language. At least the spoken form presupposes fluency and immediate reactions in interaction with someone else while the written form generally enables people to take some time if they cannot either grasp the meaning of a text or recall an appropriate word in writing.

Interaction refers to people's capability of negotiating meaning in interaction with someone else. Language use, which is a part of the linguistic field, focuses on language itself, such as grammatical correctness or different language plans, such as phonetics and phonology, morphology, syntax etc. For example, the native speakers of English are familiar with the correct usage of articles as well as the native speakers of Czech or Russian do not make serious errors in the flection of nouns.

As far as the communication of beliefs is concerned, it might be assumed that there are so many languages as there are people. The reason is that even a person who is a native speaker, let us say, of the English language, stores slightly different vocabulary in the mental lexicon

from other English people. None of them have experienced the same situations. Nonetheless, their mutual language is profoundly influenced by their mutual culture (Condor & Antaki 1997:320-47).

To define the term discourse, as Dijk (1997:25) proposes, non-verbal communication is worth taking into account. People often use gestures, mimes or facial expressions to stress what they just say. With regard to second language learners, however, they might gesture for reasons that they do not know, for example, a collocate related to *a goal keeper,* namely instead of *goal,* in the sentence *A goal keeper must stay in a goal,* these learners might use their hands to draw a frame in the air. This lack of knowledge might cause some problems in communication especially with another non-native speaker.

As far as collocations in terms of second language acquisition are concerned, everything mentioned above is highly relevant, when the fact that language knowledge is collocational knowledge is taken into account. Second language learners are forced to struggle with the fact that "most of the words in daily use have several meanings, and any occurrence of the word could signal any one of the meanings" (Sinclair 1991:103). Moreover, as Lightbown and Spada (1999:101) remark, they should attempt to be familiar with different social situations in which every single collocational expression generally occurs.

With reference to the 'binomial' (McCarthy & O'Dell 2001:164) *odds and ends*, learners ought to be taught not only its meaning but also whether it is treated as formal/informal or approving/disapproving. The reason is, for example, that the Czech translation of *odds and ends* might be used in all four situations. Similarly, some similes are formal in Czech and some informal, however, in English, they are generally treated as informal.

2.4.1 Second Language Discourse

Lightbown and Spada (1999:74) deal with Selinker's (1972) term interlanguage to name a system of language that second language learners are somehow developing as they acquire a second language. The development of second language acquisition (SLA) can be influenced by either other languages learnt before or some similar characteristics in their first language. In case of this research, the English language refers to a second language.

As far as collocations in terms of SLA are concerned, interlanguage indicates the development of learners' collocational knowledge. The purpose of developing such a system is described by Ellis (1997:122):

> Learning to understand a language involves parsing the speech stream into chunks which reliably mark meaning. Learners don't care about linguists' analyses of language. They don't care about grammar or whether words or morphemes are the atomic units of language. From a functional perspective, the role of language is to communicate meanings, and the learner wants to acquire the label-meaning relations. (Ellis 1997:122)

From the point of view of 'competence' (Chomsky 2006:102), second language learners parse the speech in order to answer the question whether it is grammatically correct. As for 'performance', the learners' memory, for example, is one of the most crucial features of a second language acquisition. The learners' memory is further influenced by the amount of exposure, for instance. If a learner of the English language does not remember the word *child* in the sentence *My father is a big child,* then he or she is not able to understand neither the literal nor the metaphorical meaning.

According to Chomsky (2006:103), linguistic competence could be limited by linguistic performance. With regard to content and function words, it might be assumed that the second language learners' competence is much less limited by the performance when these learners are familiar with all the content words but not with the function words included in the sentence *My father is a big child*. The reason is, as Ellis (1997) remarks, that they hardly care about grammaticality of language. Learners intend to communicate meanings which are carried particularly with content words.

Chunking is a natural way that is applied even by native speakers, in the case of this research the English. However, native speakers of English do not apply chunking to analyse language as linguists do. They need it to negotiate for meanings. It follows that not only learners but also native speakers use language particularly to interact or communicate with other people.

From the teaching and learning point of view, collocations "which are to some degree unpredictable, need to be examined for any patterning that occurs" (Nation 2001:335). For example, the words *head* and *tail* are frequently used in English idioms (*from head to toe*, *nose to tail*) to express some space between two opposite sides or emphasize the difference

between those sides. *Head* and *tail* do not occur with the words *heel* or *knee,* as *head* does in the Czech language.

Sternberg (1987), described in Ellis's study (1997:135), defines three basic subprocesses which take place when learners attempt to acquire a second language. The first one, called 'selective encoding', refers to the ability to distinguish between relevant and irrelevant information. Afterwards, when learners know what information is relevant, they combine it "into a workable definition" (Ellis 1997:135). The third one 'selective comparison' means that learners are able to make connection between new and old information. Furthermore, on the basis of the previous knowledge learners should be encouraged to deduce the meaning and the form of new vocabulary items.

On the other hand, as far as collocations in terms of SLA are concerned, Lightbown and Spada (1999) emphasize that "most learners believe that idiomatic or metaphorical expressions cannot simply be translated word for word" (Lightbown & Spada 1999:79). For example, a learner of English whose mother tongue is Czech probably presupposes that *as white as a wall* is possible in his or her L1 but not in English (L2).

2.4.2 Hypothesis of the study with regard to gender

This study does not intend to reveal some significant differences between how women and men acquire collocations of an L2 (in case of this research English). It is assumed that the four women might concentrate on the English language itself (in terms of collocations) more than the two men. The reason for such a statement follows from the findings of Sunderland's study (1996), described by Litosseliti (2006:82), which points out two important features.

On the one hand the teacher interacts more with boys during lessons. This interaction, however, does not comprise any so called academic exchanges, which means that the teacher addresses boys in order to tell them to behave themselves. They often have problems particularly with discipline. On the other hand, although girls interact less with the teacher than boys during lessons, girls seem to be more academically involved. To be more specific, girls hardly have serious problems with discipline and thus they can concentrate more on learning (Litosseliti 2006:82). It follows that women might be accustomed to focusing on

language itself more in detail than men, because they have been applying this strategy since childhood.

3 Interview Analysis and Discussion

This section focuses on the detailed analysis of each test and findings based on the interviews. There are three parts. Both the interview parts based on the tests attempt to describe some essential features of every single collocation which is incorporated in a sentence example. Additionally, both the parts also intend to summarize what all the participants mentioned about their own experiences, feelings of acquiring English collocations. The third part consists of the analyses of other collocations which are included in the interview but not in the tests. Finally, both the parts comprise some key findings that can help with fulfilling the aim of this research.

3.1 Multiple-Choice Test

The multiple-choice test (see appendix A) is analysed sentence by sentence, while there is a total of five sentences. Every single sentence consists of one type of collocations such as an idiomatic expression represented by *couldn't make head or tail of.* One example of content word co-occurrences *brewed beer* and one sample of content-function word co-occurrences *look at* are included in the second and third sentence.

The fourth one should represent not only a sample of content word co-occurrences but also the case when a literal meaning is substituted by a transferred meaning (*heavy traffic*). The last sample, comprised of direct speech, embodies a link between concepts, which is *a shred of*. It is also important to remark that other options from the multiple-choice test are worth noticing and analysing since they played a key role in the interviews.

This kind of test generally focuses on receptive skills such as reading and writing. Nevertheless, as the test is just one among other parts of the whole interview, the participants were asked to apply some other options in English utterances. Hence, the multiple-choice test of this research might be treated as a linguistic construct which concentrates on both productive (speaking, writing) as well as receptive skills (listening, reading).

3.1.1 Couldn't Make Head or Tail of it

The first sentence includes a type of English idiomatic language that can be classified, with regard to McCarthy and O'Dell (2001:164), as a binomial, as there are two content words *tail* and *head* linked by the function word *or.* Additionally, their co-occurrence is based on euphony (the pleasant rhythm). However, this binomial is only a part of the whole and more expanded expression *couldn't/cannot make head or tail of something,* which is fixed and no words can be replaced.

Oxford Advanced Learner's Dictionary (2005) also deals with a slightly different form of this idiom, namely *can't make head nor 'tail of sth*. In relation to Nation's classification (2001: 329-32), it might be a sample of both grammatically and lexically fossilized collocation. As for the participants' knowledge of such an idiomatic expression, none of them would include it in their productive vocabulary and only one woman had encountered this idiom before.

Even though there was only one participant who knew receptively the meaning, use and form of the collocation, the five other learners were not far from the right answer. Moreover, they did not have problems with guessing the meaning of this idiom from the context. Two of them (a man and a woman) compared the English idiom with some similar idioms in their mother tongues (Czech, German), which is in concordance with Lightbown and Spada's study (1999:79), and chose the options *head or knee of* and *head or heel of.* Yet, they felt that it was not correct in terms of the content words *knee* and *heel.* Two participants made the decision on the basis of rhythm, which paid off, since their replies were correct.

This phrase contains seven chunks of information in the form of three function and four content words. According to Miller's study (Nation 2001:319), this number is the limit for human short-term memory. Therefore, four out of the six participants found much easier to learn only the part with the binomial *head or tail* than the whole expression with the seven words. Moreover, five participants did not hesitate when they were asked to decide on the right preposition following the content word *tail,* which might be caused by the fact that the combinations of function words are restricted by the type of the test. Only one participant hesitated between *head and tail in* and *head or tail of.*

This participant justified the hesitation by explaining that she did not concentrate on such words as *and, or* (conjunctions), *of, with, in* (prepositions) at first. However, she sought a word which might have been connected to *head* either by means of a sound pattern or on the basis of whether there were near-synonyms or antonyms or logical opposites among the options. Having chosen the word *tail,* the participant was not able to deduce, as the others were, which conjunction (*and, or*) and preposition (*in, of*) belong to this fixed expression.

As far as the function words of this collocation are concerned, four out of the six participants made the right decision because they analysed it at first and then followed an English grammatical pattern. To be more specific, they first analysed the utterance without the negative *not*, which means that the sentence would have to have such a form as *I could make neither head nor tail of it.* However, these participants noticed that the modal verb *could* is in the negative form, thus they did not hesitate to make a decision about which of three combinations (*or...of, and...in, and...with*) is the correct one.

On the other hand, the grammaticality of this collocation was not recognized by the woman who did not tackle the problem of finding the second content word of the phrase. The reason was, as she remarked, that she had already encountered this pair of words *head* and *tail* in other English idioms, namely *from head to toe* and *nose to tail*. Therefore, she found it difficult to connect such content words with other function words and to use them in a different context.

Overall, everyone was able to guess the meaning of this collocation on the basis of the context in which it occurred. If the participants encountered such a phrase in a cloze test (without options), they would not be able to write the correct answer since none of them keep it as a part of their productive vocabulary. All the participants have agreed on the fact that when they bump into some fossilized expressions, either grammatically or lexically, for the first time, they attempt to acquire them from smaller chunks which they subsequently group into larger chunks as Nation (2001:319) describes. Furthermore, everyone who considered reading this idiom aloud found it really helpful in order to grasp its meaning and to decide on the right form. It follows that the rhythm plays a key role in acquiring this idiom, which is emphasized by McCarthy and O'Dell (2001:164).

3.1.2 Brewed Beer is Somewhere between Cooked and Boiled Beer

The collocation *brewed beer* in the second sentence represents an example of 'the possible range of collocates' (Nation 2001:56) and a common word co-occurrence in terms of type. Considering type a content word collocates with a content word. As to word classes, this collocation consists of an adjective and a noun. As for the possible range of collocates, if the participle adjective *brewed* is *labelled* as a node, with reference to Müller (2008:4), its frequent and common collocates are *tea, coffee* and *beer*. This means that there are only three possible noun concepts which are connected to the participle adjective *brewed*. On the other hand, as for the noun *beer*, OCD (2002) deals with twenty four possible adjectives that are commonly or frequently used together with *beer*, for example, *strong, light, foaming, blonde, draught, cask* etc.

The collocation *brewed beer* often occurs in magazines (Davies 2004) and particularly on the labels of beer products. The media exposure is regular and massive and thus all the participants had no problem to choose the right answer. However, three of them whose mother tongue is Slavic remarked identically that there is not one special word like in English to name such a process as making beer. One participant, a woman from Russia, explained that *brewed* is somewhere between *cooked* and *boiled*. There is only one word that can be used in different contexts with different meanings. These three people did not find it problematic to remember a new word *brewed,* since they knew its meaning and use. The only feature that they had to remember was a new form. According to Cruse (2004), they did not have to create a new concept in the mental lexicon.

Having found out that all the participants were familiar with the collocation *brewed beer*, it was desirable to elicit from the participants how far they knew such a limited word *brewed* in terms of 'collocates' (Müller 2008:6). Nobody were able to use it with another word nor in a different context. Everyone felt surprised by the fact that *tea* as well as *coffee* can *be brewed* too. They had never encountered such collocations as *brewed tea* and *brewed coffee* before although, according to COCA (Davies 2004), they belong to the highly frequent co-occurrence. Furthermore, both are more frequent than the expression *brewed beer*.

As the interviews followed, the participants were asked whether they were able to use the other four words from the table under the second sentence. All of them created two simple sentences in which they incorporated *boiled* and *cooked* out of the blue. *Stewed* caused some difficulty, however, surprisingly, there was nobody who could say that they had encountered the word *simmered* before. The table 3.1 depicts the relative frequency and register in which these words have occurred most often since 1990. The data are taken from COCA (Davies 2004):

Table 3.1: The relative frequency and register of simmer, boil, cook, stew.

Lemma	**Frequency (totality)**	**Register (most frequent occurrence)**
simmer	3611	magazine
boil	9045	magazine
cook	39938	magazine
stew	3087	newspaper, fiction

These figures, taken from COCA (Davies 2004), probably explain why all the participants were not familiar with the word *simmered* and found it difficult to use *stewed* in context. The reason may be that, considering Kjellmer's terminology (Nation 2001:329), their relative frequency is more than eight times less frequent than *cook*.

To sum up this chapter, the regular exposure seems to be one of the most crucial aspects in acquiring the collocation *brewed beer*, which is presented as the least frequent word co-occurrence in comparison to its collocates and the other four words from the multiple-choice test. There was no need to make a new concept in the mental lexicon when the learners acquired the words *brewed, cooked, boiled.* The only thing they had to do was to memorize a different word for a well-established concept.

From a first language acquisition point of view, the participants had succeeded in “three different but related tasks: a labelling task, a packaging task and a network-building task”(Aitchison 2003:189). Therefore, the participants might have redone a labelling task

since the only thing they probably had to change was the name of a concept which had been placed among and compared with other concepts before in their mother tongues. The cognitive process comparison is related to a packaging task. Additionally, they had already made connections between such a concept and other concepts, with regard to a network-building task.

3.1.3 Look me at Sounds Weird

According to OALD (2005), *look at* is classified as a kind of phrasal verb. On the other hand, to be more specific Quirk *et. al* (1990:338) deals with the term 'a type I prepositional verb'. Such a verb comprises two components such as a lexical verb that is followed by a preposition "with which it is semantically and/or syntactically associated" (Quirk *et. al* 1990:338). It follows that the direct object or prepositional object must be placed after the preposition, since this is the semantic core of the phrase. *Look* cannot be separated from *at* by an object as it is possible in the case of *look up,* for example. Considering a common word co-occurrence in terms of type (Nation 2001:56), a content word collocates with a function word.

When the participants heard or read the word phrase *look me at*, they all labelled it as strange or weird. Subsequently, when they were asked to explain the certainty that the phrase was incorrect, four participants (three women and one man) were able to classify *look at* as the example of a phrasal verb that belongs to a different category than *look up,* despite the fact that the forms look similar. Moreover, these four participants remembered the school teachers' highlighting that there are two categories of transitive phrasal verbs and whether they are separable cannot be deduced. Therefore, they had to learn it by heart. Two participants (a woman and a man) just felt that there was something wrong with the phrase *look me at,* yet they could not explicitly give the reason for such a certainty.

Table 3.2 shows frequency of the collocation *look at*. Additionally, there are five different kinds of register taken into account. The data are taken from COCA (Davies 2004):

Table 3.2: The absolute frequency and register of *look at*

Look at (Corpus of Contemporary American English)				
SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
54225	51565	17054	16283	8381

Obviously, this type of collocation is most frequent in spoken English. With regard to Kjellerman's term (Nation 2001:329), the absolute frequency of *look at* is roughly the same in speaking as in fiction. On the other hand, in the academic register the usage of this collocation is eight times lower than in the spoken register. From magazines and newspapers point of view, the absolute frequency of *look at* is similar to each other but still more than three and a half times lower usage compared to the highest frequent co-occurrence in terms of register.

Considering the participants' knowledge of the collocation *look at*, everyone promptly chose the correct answer in the multiple-choice test. They identically agreed that this was the most common expression used by the teachers of English at school. All the participants also agreed that they had known it for a long time since they had started acquiring English. This finding is in accordance with the figures (Davies 2004) in table 3.2 dealing with the absolute frequency of *look at*. Therefore, this phrase can be classified as a highly frequent collocation used by both native speakers as well as second language learners.

Two women and one man, whose mother tongue is a Slavic language, attempted to acquire *look at* by analysing every single item in terms of meaning. Having analysed *look* and *at* separately, they put them together and used it in speech or writing. On the other hand, the other half of the participants followed the reverse method. They first tended to remember the whole phrase together with a direct object, for instance *look at me/the picture*. Afterwards, they analysed it item by item to see how those items behaved if they were connected to other collocates. According to Nation (2001:319), some learners prefer to acquire a second language by chunking it from the smallest parts to the largest parts of speech. Some may prefer the reverse direction.

One participant emphasized: “It is much easier to remember *look at* than *look down on*.” This utterance was followed by the concrete explanation that two items together could be remembered easier and could be retrieved much faster than three or more. Afterwards, she

used both *look at* and *look down on* in two simple sentences such as *I looked at you*, *I looked down on you* and counted that there are four words in the first sentence and five words in the second one. Thus, in fact she had to acquire four words in the first case and five words in the second case. Apparently, she unconsciously realized, as Ellis (2001:39) remarks, that there exists a limited range of words or information humans can keep in the short-term memory.

To summarize this chapter, the collocation *look at* is frequently used especially in the spoken English language. All the participants were able to find the correct answer with no problems and in a few seconds, which might be caused by a great amount of exposure. Half of them preferred to acquire such phrases by following the direction 'from the smallest part of speech to the largest part of speech'; however, three others found the reverse direction 'from the largest part of speech to the smallest part of speech' more useful.

3.1.4 Heavy Does not Always Mean Heavy

The collocation *heavy traffic* can be treated either as a content word collocating with a content word according to Nation's (2001:56) co-occurrence in type or as a metaphorical expression, considering White's (1996:17) terms, which comprise a 'primary vocabulary' and a 'secondary vocabulary'. To be more specific, the word *traffic* carries 'literal meaning' and *heavy* bears 'transferred meaning', to use the terminology from Jackson and Amvela's study (2000:59). OALD (2005) deals with five more or less similar entries which describe the word *traffic*. On the other hand, the word *heavy* relates to almost four times more definitions than its collocate, *traffic*. Moreover, *heavy* occurs in some idiomatic expressions, such as *heavy hand, get heavy, make heavy weather of sth.,* etc.

When the participants were asked whether they were able to explain the differences between *heavy traffic* and *heavy suitcase*, *heavy drinking* and *heavy work*, nobody hesitated as to the explanation of the different meanings of *heavy suitcas*e and *heavy traffic*. They knew that in the first case *heavy* refers to weighing a lot 'in terms of kilograms' (literally said by one participant), which is the first definition presented in OALD (2005). Every participant used the word *kilograms* to define the collocation *heavy suitcase*. In contrast to *heavy suitcase*, the participants explained the collocation *heavy traffic* as the traffic which is 'busy with cars' or

by saying that there are 'lots of cars together'. Besides, they all felt that *heavy traffic* has almost nothing in common with *heavy suitcase*.

The difference between another two collocations, *heavy drinking* and *heavy work,* might be treated as similar to the difference between those explained above *heavy traffic* and *heavy suitcase. Heavy* connected to *drinking* carries more or less a 'transferred meaning' since the meaning of such a collocation refers to the act of drinking a large amount of alcohol regularly. *Heavy drinking* has no relation to kilograms. However, when it is linked with *work*, we get back to the meaning of kilograms or weight because the expression *heavy work* refers to an act which is physically exhausting. It is assumed that the receptively comprehensive knowledge of the word *heavy* helped the participants infer all the differences between these four word co-occurrences in terms of meaning. (Receptively because almost everyone chose a wrong word in the multiple-choice test.)

According to the Corpus of Contemporary American English (Davies 2004), the absolute frequency of *heavy traffic* is 201 and it shows quite a high score of its semantic bonding, which is 12.48. “Typically, scores of about 3.0 or above shows [sic] a 'semantic bonding' between the two words.” (Davies2004). The meaning of *heavy* is closely dependent on the word *traffic* and vice versa. Yet, the relative frequency of *heavy traffic* is much higher.

Table 3.3 deals with the relative frequency of the collocation *heavy traffic.*

Lemma	**Frequency (totality)**
HEAVY	33795
TRAFFIC	20594

Obviously, the relative frequency of *heavy* is higher than the relative frequency of *traffic*, which is in concordance with a variety of uses defined in OALD (2005). The frequency of co-occurrence of *heavy traffic* is more than a hundred times lower than the separated words *heavy* and *traffic.* It follows that “collocations are necessarily less frequent than their constituent collocates, and items which intuitively seem to be collocations often have a very low-frequency of occurrence” (Nation 2001:333).

With regard to the relative frequency, it might be one of the crucial facts why four participants did not know the correct answer and two of them did not feel confident about it in the multiple-choice test. Even though they were familiar with all the options (adjectives), they felt confused when they were trying to match the word *traffic* with its collocates. To explain the confusion, they all agreed that they were accustomed to acquiring such expressions separately word by word. Furthermore, the teachers never taught them these word co-occurrences intentionally together, which might have profoundly influenced their learning style.

To sum up, it is obvious that if the learners of English are familiar with both the words *traffic* and *heavy* separately, they are somehow capable of guessing their meanings from a context. However, this knowledge does not guarantee that second language learners are able to make a connection between these two collocates productively. Since the score of semantic bonding of *heavy traffic* ranges more than the four times higher number than the limit is, this collocation, as Nation (2001:335) highlights, had better be acquired not exactly as one word but by examining as many patterning as possible. All the participants acquire such, to some degree, an unpredictable collocation separately, which sometimes causes some misunderstandings in communication.

3.1.5 A shred of Evidence

A shred of, a sample of a link according to Cruse's study (2004:126), is included in the last sentence of the multiple-choice test. The link *a shred of* is of the specific kind. It is used in connection with some abstract nouns that cannot be counted such as *a shred of hope*, *a shred of evidence* etc. With reference to Harley (2006), the link *a shred of* contains two function words (one indefinite article *a* and one preposition *of*) and one content word *shred*. While the preposition *of* cannot be replaced with another preposition, the indefinite article can be replaced by other function words such as *every, the, each* etc.

The word *shred* carries either the literal or the transferred meaning (Jackson & Amvela 2000:59). As for the 'literal meaning', OALD (2005) deals with the definition that *shred* as a noun (it can also occur as a verb) refers to “a small thin piece that has been torn or cut from sth” (OALD 2005). When *shred* is used as a 'quantifier' with a preposition *of* in a positive utterance, then it generally bears the fixed or literal meaning. Moreover, it is possible to have

more than *one, two, three shreds of paper.* However, in case of negative statements such as *There isn't a shred of evidence* or *Hardly can I find a shred of evidence* then this expression is commonly used only in the singular. Additionally, the plural form of *shred* as a quantifier frequently occurs in relation to uncountable or mass nouns that can be treated as concrete with literal meanings.

On the other hand, the singular *shred* in connection to abstract nouns such as *evidence* or *hope* does not really attempt to cut those phenomena into pieces. As Sinclair (1991) genuinely remarks, "some semantic feature of N2 is pointed up by the metaphor of N1" (Sinclair 1991:90). In case of the phrase *a shred of evidence,* N2, represented by *evidence,* mirrors semantically in N1 *shred.* To be specific, an utterance such as *there isn't a shred of evidence* refers to that there is neither a small nor a thin piece of this phenomenon.

According to Oxford Collocations Dictionary (2002), the utterance *there is not a shred of evidence* is treated as one phrase. The expression *a shred of* does not occur in the section of quantifiers among, for example, *a piece of evidence, a mass of evidence* and *a body of evidence.* The table 3.4 depicts the absolute frequency of every single link collocated with the concept *evidence*. Furthermore, there are also different kinds of register taken into account. The data are taken from COCA (Davies 2004):

Table 3.4: The absolute frequency and register of four links with the concept *evidence*

Lemma	**Frequency (totality)**	**Register (most frequent co-occurrence)**
shred of evidence	130	spoken
piece of evidence	512	spoken
mass of evidence	9	spoken, magazine
body of evidence	260	academic

Even though the expression *shred of evidence* more or less belongs to idiomatic language, its range of absolute frequency is only four times lower than the common word co-occurrence *piece of evidence.* Since s*hred of evidence* frequently occurs in the media based on coverage,

then it might be assumed that this is the reason why two participants who were interested in news immediately chose the correct link among other options. With regard to gender, it was the men who knew this expression.

Five participants agreed that they would never learn such an expression, *a shred of evidence*, as one word. Similar to the collocation *heavy traffic,* they were used to separating *shred* from *evidence* when they attempted to acquire it consciously. They usually searched for the literal meaning of *shred*. Afterwards, when the participants had found its meaning particularly in a bilingual dictionary, they did not further care about whether it frequently occurred in the negative utterances connected to the abstract and uncountable noun *evidence*.

Only one participant (a woman) was accustomed to acquiring such English idiomatic expressions, *a shred of evidence* or *a few nuggets of information* etc., as one word since she had been learning another foreign language in this way when she started to learn English. Additionally, this woman felt that when she kept this way, her fluency was getting better. On the other hand, she realized that the older she grew and the more advanced level she attained, the more she tended to analyse every foreign language that she acquired.

Overall, the collocation *not a shred of evidence* had better be treated as an idiom, therefore it should not be acquired separately. Most of the participants do not follow this way, since they either do not concentrate on these links or they would never say that *shred of evidence* is somehow lexically and grammatically fossilized (Nation 2001:329-32). Their statements, apart from one woman, is in accordance with what Ellis remarks (1997:122) on the fact that second language learners do not care about some grammatical features, everything they intend to do with their knowledge of L2 (in this case English) is to communicate or negotiate for the meaning.

3.2 The Key-word Transformation Test

Similar to the multiple-choice test, the key-word transformation test (see appendix B) is used as a part of the interviews. It consists of five sentences, which are analysed collocation by collocation. Such a test particularly focuses on productive skills (speaking and writing), especially on writing. Yet, as the participants attempted to rewrite the utterances in order that

they have similar meanings to the sentences given, they were interviewed by the researcher to discuss their answers. Hence, every participant often spoke more than wrote replies.

Both the first and the second sentence contain samples of the English expressions in which a content word collocates with a function word. Nevertheless, there are some differences between these two collocations. Although both are of the same type according to Nation's classification (2001:329-32), from a linguistic point of view, according Dijk (1997:2), one of them is a sample of a phrasal verb, namely *put through.* The other one embodies a lexical word that should be followed by a direct object and a preposition that is frequently used with this lexical word, the expression is '*accuse sb of sth/doing sth*' (OALD 2005).

The third sample of a collocation represents the case when a content word collocates with a content word, *had better.* The fourth sentence comprises an example of a simile such, *as white as a sheet.* The fifth and the last one exemplifies a grammatically fossilized collocation, namely *I've got that song on the brain.*

3.2.1 Accuse of is less Common than Blame for

From the psychological point of view, as Harley's study (2006:186) describes, *accuse* as a lexical word bears the meaning and *of* can be treated as a link between two different concepts. Linguistically, *accuse of* belongs, as Quirk *et. al* (1990:341) suggest, to the second type of prepositional verbs. In comparison with *look at,* where the verb *look* can be followed by an object just after the preposition *at*, the verb *accuse* can be separated by a direct object from its fossilized preposition *of*. Specifically, while it is impossible to say *look me at*, *accuse me of* is both semantically and grammatically feasible.

There were two participants who found difficult to put the verb, preposition and direct object in the right order. They intended to swap the positions of the preposition *of* and a direct object. The other four participants had no problem with putting them in the right order. Two participants were familiar with the meaning of this prepositional verb. However, only one was able to fill the blank with both the verb *accused* and the right preposition *of* without searching the meaning in a dictionary.

The preposition *for* was given to the participants instead of the lexical word *accused* in an attempt to investigate if they found it easier to recall a content word with the similar meaning to *accuse* that is generally followed by the function word *for*. None of them were able to complete the task, nor the participants who were familiar with the meaning of *accused*. All the participants agreed that it was much more difficult to recall a lexical word frequently followed by the function word *for*. Yet, when the participants heard the word *blame*, everyone emphasized that the verb *blame* was more frequent than *accuse*. Moreover, they marked this verb (*blame*) as the one they had used before at least once in writing or speaking.

The participants had never searched for the frequency of these two collocations, as they admitted. Their general feeling based only on their experiences is unfortunately not in concordance with COCA (Davies 2004). Obviously, according to the figures from COCA (Davies 2004), *accuse of* is twice more frequent than *blame for*. The main area where the prepositional verb *accuse of* occurs most is in the media focused on coverage, which is similar to the expression *a shred of evidence*. Table 3.5 deals with the absolute frequency of two prepositional verbs. Furthermore, a category of register is included. The data are taken from COCA (Davies 2004):

Table 3.5: The absolute frequency and register of *blame for* and *accuse of*

Lemma	**Frequency (totality)**	**Register (most frequent co-occurrence)**
blame (sb) for	3571	newspaper
accuse (sb) of	7097	spoken, newspaper

Four participants (women) attempted to acquire such prepositional verbs together with their appropriate prepositions. On the other hand, everyone felt that what was most important in that case was whether they knew the meaning of the content words *accuse* or *blame*. It is assumed that if the participants need to use this verb in communication, they are sometimes influenced by their mother tongues. The reason is that the participants just linked a direct object with a prepositional object by using an English preposition but semantically similar to their L1, although they were convinced that it must have been incorrect. For example, there

appeared such expressions as *accuse for*, *accuse on* or *accuse from*. This strategy is in concordance with Lightbown and Spada's hypothesis (1999:79) that learners generally feel that they cannot translate most of the idioms of an L2 word for word.

To summarize this chapter, *blame for* is less frequent than *accuse of.* The participants assumed that *blame for* was much more frequent than *accuse of* because they often encountered this phrase at school. Moreover, three of them never met the verb *accuse* before. All the participants agreed that knowing the meaning of content words enabled them to communicate somehow. In relation to Dijk (1997:25), their 'interlanguage' (Lightbown & Spada 1999:74) was often accompanied with non-verbal language such as gestures and facial expressions. The lack of knowledge of appropriate function words often interfered their fluency. On the other hand, they had never met with any critical reaction when they applied their strategy in interaction with an English native speaker.

3.2.2 Put me Through to the Manager

In the second utterance the verb *connect* is supposed to be replaced with the phrasal verb *put through.* From the psychological point of view, it is a sample of the co-occurrence in type when Nation's division (2001:56) is taken into account, namely a content word collocates with a function word. However, single *put* does not convey the meaning of *connect*. Without the preposition *through* the fuzziness of *put* becomes a dense fog. With regard to Aitchison (2003:152), both *put* and *through* must be classified as bricks that can be connected to other bricks by applying some mortar. To be more specific, when the expression *put through to the manager* is analysed, then *put* and *through* can convey the specific meaning of connect only together and thus this expression can be treated as one 'concept' (Cruse 2004) labelled by a two-content-word expression. Therefore, there are, in fact, two content words, namely *put through* and *manager*, connected with each other by means of one function word *to.* Additionally, as Harley (2006:212) remarks, *put* as a content word accommodates its meaning to the function word *through*.

As far as linguistics is concerned, *put through* belongs to the category of phrasal verbs. Namely, such phrasal verbs that are transitive and separable. It is transitive because it “may take a direct object” (Quirk *et al.* 1990:337) and separable since the direct object should be

placed between *put* and *through* when it is a pronoun, a function word such as *me, him, her, us* etc. On the other hand, when the direct object is represented with a function word, the English pattern allows speakers to make a choice, namely *put the man through* or *put through the man*. It is possible to place the object either between the verb and its particle (fossilized preposition whose value is roughly the same as the value of a verb in terms of meaning) or just after the whole expression.

The collocation *put through*, however, causes other difficulties for second language learners, especially when such kinds of collocations do not exist in the system of their mother tongues. OALD (2005) describes three examples of the general usage of *put through.* In fact it might be treated as one word with three different meanings. These meanings are not deducible from every single item of the collocation.

Moreover, second language learners ought to be careful about the animity of a direct object because this also influences the meaning of *put through*. Thus, when *put through* is used for expressing the act of connecting, the particle *through* is followed by the preposition *to*. The function word *to*, in turn, is further followed by a prepositional object, *the manager*, as in the sentences from the key-word transformation test *The operator put me through to the manager.*

Table 3.6 deals with the absolute frequency of the phrasal verb *put through.* The totality of this expression embodies the total of both the cases of co-occurrences, namely with a direct object between and just after the particle. Additionally, there is also the relative frequency of *put* and *through* depicted. The data are taken from COCA (Davies 2004):

Table 3.6: The absolute and relative frequency of the phrasal verb *put through*

Lemma	**Frequency (totality)**	**Register (the most frequent co-occurrence)**
put	58719	spoken, fiction
through	97712	fiction, academic
put (sth/sb) through	497	magazine, spoken
put (sth/sb) through to (sb)	18	fiction

The figures in the table proves Nation's statement (2001) that the bigger chunks are, the less frequently they occur. All the participants did not learn phrasal verbs separately, namely they focused at first on the meaning of *put* and then *through*. Three of them (one man and two women) acquired such verbs together with a direct object between the verb and its particle. The other half just knew that *put through* conveys different meaning than *put up with*. Whether it is separable, three participants did not find it such an important and crucial fact that could cause some serious problems in communication.

Five participants were familiar with both the form and the meaning of the phrasal verb *put through*. They identically said that such an expression belonged to their productive knowledge of English vocabulary. They often needed to use it either in everyday life or at university since they had to make a great number of phone calls in English per month. Only one participant, the youngest one, had not encountered this phrasal verb before, neither at school nor in daily life.

Overall the phrasal verb *put through* must be acquired as one word like its synonym *connect*. Moreover, it may enhance learners' fluency when they learn transitive and separable phrasal verbs with a direct object placed between the verb and its particle, since only the knowledge of the meaning and form of such collocations does not seem to be sufficient. The use is in this case also a crucial feature. All the participants in this research were used to acquiring phrasal verbs like one word.

3.2.3 Had Better

The third example of collocational expression may be treated, with regard to Nation (2001:329-32), as either grammatically and lexically fossilized or a word co-occurrence in type where a content word collocates with a content word. However, we can also analyse it as the case in which a function word collocates with a function word. The reason is that both the words *had* and *better* in this collocation do not behave like, with reference to Aitchison's terms (2003:152), two independent bricks which should be connected with some mortar. It is the similar case to the phrasal verb *put through*. They rather function as a link between two concepts.

For example, considering Cruse's study (2004:185-216), the sentence from the key-word transformation test *Pavel had better go to the dentist., Pavel* is one concept and then there are two others *go* and *dentist.* To link the concept *Pavel* to the concept *go*, the 'function collocation' *had better* is used. The synonym of *had better*, which is *should,* is less emphatic, while *had better* may rather express some "urgency, warning and threat" (Foley & Hall 2003:193).

Table 3.7 compares the frequency of two words with a similar meaning. In addition, there is also the absolute frequency of the contracted form *'d better*. The one-word expression *should* is apparently of the highest frequency particularly in spoken language. The contracted form *'d better* is more than twice frequent than the non-contracted form *had better*. The figures prove Nation's remark (2001:321) that bigger chunks refer to lower frequency. It might be assumed, on the basis of these figures, that the frequency plays an important role in the finding that the participants were familiar with the contracted *'d better* and not with the non-contructed form *had better*.

Table 3.7: The frequency of three different word expressions but only in terms of form, since they have similar meaning.

Lemma	**Frequency (totality)**	**Register (the most frequent co-occurrence)**
should	299351	spoken, academic
had better	1682	fiction
'd better	3870	fiction

There was no participant who was familiar with the form of this collocation. Five of them were convinced that they had already encountered *better* with another word, which had a similar meaning to *should*. Three participants used a contracted form, namely *Pavel'd better go.* However, even these participants did not know the non-contracted form *had*. One woman compared this expression with those which she was already familiar with. On the basis of her knowledge of English vocabulary, she presupposed that the contracted form might have been *would* like in another word expression *would rather.*

In relation to Sternberg's subprocesses (Ellis 1997:135), she compared her previous knowledge of English with new information mainly in terms of productive skills. Afterwards, this woman attempted to infer the correct form. Unfortunately, the collocation *had better* cannot be deduced on the basis of *would rather*. It belongs to the category of word co-occurrences that contains "some element of grammatical or lexical unpredictability or inflexibility" (Nation 2001:324). In this case selective comparison is not reliable.

Overall the collocation *had better* is to be acquired together, since it behaves like a one-word expression. Additionally, its non-contracted form cannot be deduced on the basis of previous knowledge of English. It is much less frequent than the expression *should,* even its contracted form *'d better*. Although three participants were able to use the contracted form *'d better*, they agreed on the fact that neither *had better* nor *'d better* did not belong to their productive knowledge of English vocabulary.

3.2.4 As White as a Sheet next to Blank, Wall and Paper

The fourth example of an English collocation can be treated as a kind of metaphorical language. According to Jackson and Amvela (2000:65), it is a type of the multiword lexeme. Learners cannot predict its meaning from its components. However, in the case of 'similes', McCarthy and O'Dell's term (2001:162), sometimes it might be possible to deduce receptively the meaning of such expressions. As to the simile *as white as a sheet*, the form seems to cause more problems than its meaning.

With regard to Harley's study (2006), this idiomatic expression comprises two content words (*white* and *sheet*) and three function words (*as, as, a*). Since, according to Nation's classification (2001: 329-32), it is impossible to change somehow the order of every single word in the idiom, it might be labelled as grammatically fossilized. Moreover, we can treat it as lexically fossilized but only partially, since McCarthy and O'Dell (2001) suggest "you can usually make a simile using **as … as can be**" (McCarthy & O'Dell 2001:80). It follows that a *sheet* can be replaced by *can be*, namely *as white as can be.*

The metaphorical expression *as white as* can be also used in connection with a different content word than *a sheet*. Namely, next to the simile *as white as a sheet,* the other one *as*

white as snow exists. However, as Sinclair remarks (1991:103), any occurrence of a word is influenced by other words surrounding it. Therefore, when one participant (a woman) wrote *a sheet* at first instead of the word *snow*, she did not realize that the simile *as white as snow* conveys a different meaning in terms of pragmatics, since the expression with *snow* is usually used to describe someone who is beautifully white. On the other hand, *as white as a sheet* describes a person who is pale with fear, for example.

Table 3.8 depicts the absolute and the relative frequency of both the similes *as white as a sheet* and *as white as snow.* There is only the relative frequency of all content words such as *white, sheet, snow.* Their absolute frequency is quite low and thus McCarthy and O'Dell (2001:80) propose that learners of the English language should rather keep them as a part of their receptive or passive vocabulary. Although the absolute frequency is much lower than the relative frequency, learners should be familiar with these expressions at least receptively due to highly frequent words comprised of similes.

Table 3.8: The absolute and relative frequency of *as white as a sheet* and *as white as snow.* The data are taken from COCA (Davies2004).

Lemma	Absolute Frequency	Relative Frequency
As white as a sheet	7	x
As white as snow	12	x
white	x	175922
sheet	x	11373
snow	x	24824

Considering the participants' knowledge of this metaphorical expression, none of them were familiar with its form. They somehow deduced the meaning of an idiom that should be filled in. However, they were not able to use the correct form. Furthermore, everyone identically said that if there was not the key word *white,* they would find it insoluble. In order to finish the task, all the participants translated this idiom from their L1 into English word for word, although, as Lightbown and Spada (1999:79) remark, they were convinced that it was not a native-like expression.

Three out of the six participants likened someone whose face went pale with fear to *a wall,* because in Czech and Austria people generally say *as white as a wall*. Two participants, who hesitated between *wall* and *paper*, finally wrote the word *paper.* The woman from Russia considered that it might have been *blanket* similarly to her L1. Only one woman did not attempt to translate the idiom word for word, however she made Sternberg's 'selective comparison' (Ellis 1997:135) and tried to find the relation between old and new information. Afterwards, she used *snow* as was mentioned above.

Overall every participant did not count similes as the part of her or his productive skills. When they encountered such expressions in communication, they rather attempted to guess the meaning from the context in which they occurred. Similes belong to the collocations with low frequency. However, the frequency of every component of such idioms is high. Moreover, most English similes belong to the informal rather than the formal register (McCarthy & O'Dell 2001:162), which means that English similes are often humorous and used in some informal communication, for example, among close friends. It follows that register is one of the most crucial features that should be taken into account in terms of a second language acquisition.

3.2.5 I've got that Song on the Brain or in my Brain

The idiomatic expression *I've got that song on the brain* can be labelled, with regard to Nation's classification (2001:329-32), as grammatically fossilized since *that song* can be replaced by the concrete name of a song. For example, it is possible to say *I've got Mamma Mia on the brain.* On the other hand, the preposition *on* cannot be changed, nor the content word *brain.*

Considering Jackson and Amvela (2000:59), the whole expression has the 'transferred meaning'. According to White's terminology (1996:17), this idiom can be divided into two parts. The first one consists of *I've got that song* and might represent 'a primary vocabulary', the second part, *on the brain,* might introduce the metaphorical language or 'a secondary vocabulary'. Even second language learners promptly deduce when they bump into this phrase in communication that there is nobody who has a song on his or her brain literally.

It can be assumed on the basis of this research that second language learners can infer the meaning of this expression by applying three subprocesses, as Sternberg defines (Ellis 1997:135), when they encounter it in context. The reason is that all the participants did not find it difficult to grasp the meaning of the idiom expressed in the fifth sentence. Nevertheless, when the participants were asked to transform the fifth sentence into idiomatic language, none of them were able to finish it successfully. They felt that they had already encountered this collocation, yet they had never used it productively.

All the participants, even though they knew that they were wrong, compared this English expression to the ones in their mother tongues. In spite of the fact that not everyone shares the same L1, five participants wrote similar answers, namely *song in my brain*. Two of them added the word *stuck* between *song* and *in my brain*. Only one woman (from Spain) transformed the fifth utterance into *song under my brain.*

Apparently, all the participants used the personal pronoun *my* for determining the noun *brain*. When they were asked to explain it, half of them did not know any explanation. The other half, on the other hand, were convinced that parts of the body are generally used with personal pronouns. With regard to Sternberg (1987), described in Ellis's study (1997:135), three out of the six participants in this research consciously made 'selective comparison' and on the basis of previous knowledge of English decided on a kind of determiner.

The cognitive process comparison also took place when the participants used the wrong preposition *in.* Yet, in this case five of them attempted to find the correct answer in the system of their mother tongues although, as Lightbown and Spada (1999:79) remark, they believed that it should not be translated word for word. The woman from Spain, whose answer was *under my brain,* rather made a guess because, as she highlighted, such phrases are usually different in every language. Moreover, as she added, the English culture is totally different from the Spanish culture.

To sum up, all the participants were not familiar with the idiom *I've got that song on the brain*. Everyone agreed that they would never learn it as one word since it was comprised of many words and its meaning could be easily deduced on the basis of context. In order to use such a collocation productively, all the participants either consciously or unconsciously

sought through their previous knowledge of English for fulfilling the task. Additionally, they translated the English expression word for word although they knew it must be different.

3.3 'I want to marry you' sounds better

At first sight the utterance *I want to marry you* may not be treated as a type of collocation. However, with regard to Ellis (1997:129), non-native speakers sometimes tend to create sentences which are grammatically correct but it is difficult to grasp their meaning. The reason is that their grammatically correct expressions are out of native-like control (*I wish to be wedded to you, Your marrying me is desired by me, My becoming your spouse is what I want*).

Firstly, the sentence *I want to marry you* contains only five words compared to the three others. It is the shortest one. Additionally, while the native-like sentence comprises simply the agent in one form *I*, the sentence *Your marrying me is desired by me* consists of the subject pronoun *I* which is modified to two object pronouns, *me*. Secondly, there is a highly frequent collocation in the English sentence, namely *want to*. Considering Nation (2001:56), it is the co-occurrence in type where a content word collocates with a function word.

All the participants agreed that three non native-like utterances must be incorrect. They felt that something was wrong with them. According to four participants, they were not English. Two participants were convinced that these three sentences were not grammatically correct, which is in concordance with Chomsky's theory (2006:102). When the participants were asked to rephrase them somehow, everyone was able to do it.

Overall, even non-native speakers, who have some experiences of English, are able to recognize when an utterance sounds non native-like and foreign. The participants felt that "[w]ords do not occur at random" (Sinclair 1991:110). Additionally, they knew, similarly to their L1, that English as a language is only a part of other symbolic systems of a society (Zuengler & Miller 2006:39).

3.4 Overall findings

Finally, it might be useful to summarize the findings of this research. There are six most important findings that attempt to answer the questions from the aim of this study. It is obvious that this research has some significant weaknesses, which are described in the last paragraph of the list.

- It was found out that second language learners acquire collocations by means of chunking which moves in two directions. Some learners prefer learning collocations from smaller chunks which they group into bigger ones to learning collocations from bigger to smaller chunks. Some prefer the reverse direction.
- The rhythm as well as the regular exposure seem to be really helpful for learners to acquire idiomatic language.
- Knowing the meaning of content words enables learners to communicate, however it is not sufficient in order to enhance fluency.
- Phrasal verbs and idioms (fossilized expressions) are supposed to be acquired as a one-word expression.
- Both absolute and relative frequency are supposed to be taken into account while acquiring collocations because the bigger chunks are, the less frequently they occur.
- There were no significant differences found between how the women and the men in this research acquired collocations. The only difference which appeared was that the two men cared less about the English language itself than the four women. These two men agreed that they used English to communicate meanings and if they were understood, everything was possible.
- There are four major weaknesses of this research. The research had to be finished in quite a short time (more or less in six weeks), therefore everything was adjusted to this fact. The number of participants (only six) and the choice of two kinds of tests might influence overall findings. Additionally, only learners who had reached the intermediate level were interviewed.

All in all, hard work is hidden behind collocational knowledge. Even though every learner is a unique human being influenced by a social background from which he or she comes from to an English lesson, most second language learners of English probably acquire English collocations particularly by means of chunking, guessing, comparing. These three cognitive processes appeared to be common to all the six participants of this research.

4 Conclusion

This study attempted to investigate how second language learners acquire English collocations. Chunking is one of the most important strategies all the participants apply. Which direction they use depends on the type of collocations and lexical or grammatical unpredictability of their elements. The participants also rely on their previous knowledge of English as well as their L1. When they encounter a new vocabulary item in context, they attempt to connect old information with new information. It follows that they make comparison and guess the meaning, form and use from context.

All the participants find it more difficult to acquire a collocation in which a content word collocates with one or more function words. The rhythm is particularly helpful when learning some idiomatic expressions that are grammatically or lexically fossilized. The participants do not care about linguistic terms, they use English to communicate meaning. They feel that the English language has different rules from their L1. On the other hand, the participants are convinced that the only thing they should do to improve their English (L2) is to use it in every possible way.

5 List of References

Aitchison, Jean. (2003) *Words in the Mind: an introduction to the mental lexicon.* 3rd edition Oxford: Blackwell.

Bachman, Lyle F. & Adrian S. Palmer. (1996) *Language Testing in Practice*. Oxford: Oxford University Press.

Chomsky, Noam. (2006) *Language and Mind*. Cambridge: Cambridge University Press.

Condor, Susan & Charles Antaki. (1997) 'Social Cognition and Discourse', in Teun A. van Dijk (ed.) *Discourse as Structure and Process*. London: Sage Publications, 320-345

Cruse, Alan. (2004) 'Word meanings and concepts.' *Meaning in Language: An Introduction to Semantics and Pragmatics.* Oxford: Oxford University Press, 125-137

Davies, Mark. (2004-) *Corpus of Contemporary American English* Available online at http://www.americancorpus.org/

Dijk, Teun A. van. (1997) 'The Study of Discourse', in Teun A. van Dijk (ed.) *Discourse as Structure and Process*. London: Sage Publications, 1-34

Ellis, N. C. (2001) 'Memory for language', in P. Robinson (ed.), *Cognition and Second Language Instruction.* Cambridge: Cambridge University Press, 18-126

Ellis, N. C. (1997) 'Vocabulary acquisition, word structure, collocation, word-class, and meaning', in Schmitt and McCarthy (ed.), *Vocabulary: Description, Acquisition and Pedagogy.* Cambridge: Cambridge University Press, 122-139

Harley, Heidi. (2006) 'Lexical Semantics: The structure of Meaning, the Meaning of Structure', in H. Harley (ed.), *English words: a linguistic introduction*. Oxford: Blackwell Publishing Ltd, 185-216

Jackson, Howard & Etienne Zé Amvela. (2000) *Words, meaning, and vocabulary: an introduction to modern English lexicology*. London: Continuum International Publishing Group

Lichtman, Marilyn. (2006) *Qualitative Research in Education: A User's Guide*. London: Sage Publications.

Lightbown, Patsy M. & Nina Spada. (1999) *How Languages are Learned.* 2d Oxford: Oxford Univ. Press.

Litosseliti, Lia. (2006) *Gender and Language: Theory and Practice*. Oxford: Oxford University Press.

McCarthy, Michael & Felicity O'Dell. (2001) *English Vocabulary in Use: upper-intermediate*. Cambridge: Cambridge University Press.

Müller, Yvonne. (2008) *Collocation, Selection Restriction and the Teaching of the English Verb Grammar.* Munich: GRIN Verlag.

Nation, I. S. P. (2001) *Learning vocabulary in another language.* Cambridge: Cambridge University Press.

Oxford Advanced Learner's Dictionary. (2005) 7. uppl. Oxford: Oxford University Press.

Oxford Collocations dictionary for students of English. (2002) Oxford: Oxford University Press.

Quirk, Randolph. *et. al* (1990) *A Student's Grammar of the English Language*. England: Longman.

Sinclair, J. M. (1991) *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

White, Roger M. (1996) *The Structure of Metaphor: The Way the Language of Metaphor Works*. Oxford: Blackwell Publishing Ltd.

Zuengler, Jane & Elizabeth R. Miller. (2006) Cognitive and Sociocultural Perspectives: Two Parallel SLA Worlds?. *Tesol Quarterly* 40, 35-51.

6 Appendix A

Multiple-choice test

1. Rosie said something about keys but I couldn't make head _____ it.

a) or heel of	b) and tail in	c) or knee of	d) and heel with	e) or tail of

2. The neighbouring town of Faraway Downs has a delightful pub with its own _____ beer .

a) cooked	b) brewed	c) boiled	d) simmered	e) stewed

3. Students had to _____ the teacher and guess whom he was taking off.

a) look on	b) watch at	c) look down on	d) look at	e) see on

4. Most great cities of the world struggle with _____ traffic every day.

a) wrong	b) heavy	c) serious	d) severe	e) big

5. " The public is shocked when they learn that there isn't _____ evidence for a lot of medical therapies, " says Dr. Richard Greene. (Davies 2004)

a) a few nuggets of	b) a snippet of	c) a kind of	d) a shred of	e) a fragment of

7 Appendix B

Key-word transformation test

Complete the second sentence so that it has a similar meaning to the first sentence. You can use between two and five words, including the word given.

1. "You've stolen my mobile phone, George!" said Jane.
accused
Jane ….....having stolen her mobile phone.

2. The operator connected me to the manager.
put
The operator ….....the manager.

3. Pavel should go to the dentist.
better
Pavel.....the dentist.

4. When Malin saw the mask, her face went pale with fear.
white
When Malin saw the mask, her face went as.....

5. I cannot stop myself singing that song.
brain
I've got that.....

8 Appendix C

Interview questions

1. Have you ever heard of the term collocations?
In case of yes, could you remember the place? Was it at school or somewhere else?
2. I am going to read three sentences. Could you tell me whether they are correct and if you think that there is something wrong with them, may you explain why you think so?
I wish to be wedded to you.
Your marrying me is desired by me.
My becoming your spouse is what I want.
3. Would you ever say that *tomorrow will be* ***strong rain***? Explain why yes or no.
4. How would you interpret if your friend told you that her or his father *kick the bucket yesterday*?
5. I am going to read a sentence with a blank. Think about the word you can fill in. Here are five options from which you can choose. Was it helpful when I said it aloud? Do you know this expression *head or tail of* ? Would you say the same in your mother tongue or are there any differences? In case of yes, could you translate your mother-tongue version into English?
6. How would you say brewed beer in your mother tongue? Do you have an extra word for this expression as well as English does? What else can be also brewed? Look at other four words in the chart, please. Could you tell me what words collocate with them? Will you remember any?
7. How long, you think, have you known the word look at? Do you know other words similar to this and how do you learn them? Do you write/say them together or separately when you acquire them consciously? Would you ever say *look me at* if you usually utter *look it up*?
8. What is the difference between *heavy traffic* and a *heavy suitcase*, *heavy drinking*, *heavy work* etc.? Do you know some more collocations with the adjective heavy?
In case of a wrong answer:
Why did you answer like this? Did you compare it to your language or did you make a guess because the English say *severe traffic jam/serious accident*?
9. Have you ever heard this expression before? Do you remember a word that frequently follows the links *a fragment of, a few nuggets of, a snippet of?* When you say a shred of, will you be able to explain the meaning of the word *shred*? I am asking since I won't. I learnt it like *a piece of/ a kind of/ a type of...*
10. Which collocation is the most frequent in you spoken and written English? Additionally, which one do you only encounter but do not use on your own?
11. How do you learn such words? Do you prefer to acquire the meaning of the word accuse at first and then when you have to use it you just guess a preposition? On the other hand, do you learn it together simply like one word *accuse of* or do you remember the context in which you encounter/bump into the word? When I give you the word *for*, will you be able to use it correctly, too?
12. Why do you think a second language learner finds it difficult/hard to learn phrasal verbs?
13. Why did you write would/had? Is had better your frequently used word or you hardly encounter it?
14. Is the expression as white as a sheet similar to your mother tongue? Did it help you with your decision?
15. How would you say that you cannot stop singing a song idiomatic in your mother

tongue? I am asking because it is different in the Czech language.

16. Haw long have you been learning English?
17. Why do you think second language learners find collocations so difficult and hard to acquire? Do you feel the same? May I say *so heavy to acquire*?
18. Have you ever experienced such misunderstanding as I mentioned at the beginning of this interview with the expression kick the bucket? What happened? I have with write a test instead of take/do a test.
19. How do you generally acquire such phrases as they are in the tests? Is it helpful to go abroad to speak more fluently? Why yes/no.
20. Can you remember a word that exists in English and doesn't in your mother tongue? Are you able to use it in an English sentence? (a stile – in Czech)
21. Do you have anything you want to add that we have not talked about?