



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Depth of vocabulary knowledge of MA students of
Chinese studies“

verfasst von / submitted by

Stephanie Dalton, BA

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Arts (MA)

Wien, 2020 / Vienna, 2020

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

A 066 811

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

MA Sinologie

Betreut von / Supervisor:

Univ.-Prof. Mag. Dr. Susanne Weigelin-Schwiedrzik

Abstract

This study investigates the depth of vocabulary knowledge of MA students of Chinese Studies at the University of Vienna. In particular, the focus lies on these learners' identification of one specific type of word association, namely collocations. This study also examines how target word frequency, as well as other factors (e.g. language learning background or a stay abroad in China) impact the participants' ability to identify collocations. Furthermore, this study also compares different groups of learners among the participants (i.e. first and third semester MA students, heritage and non-heritage learners).

For this purpose a Word Associates Test was conducted, which was supplemented by a demographic background questionnaire in order to better classify the results. Data was collected from sixteen participants taking part in a language course of the MA studies programme.

The results of the study show that the participants know a large number of target words but their ability to identify collocations lags behind. Most factors mentioned above seem to have no clear impact on the participants' results.

Certain limitations of this study arise due to the small number of participants, making between-group comparisons problematic. However, useful implications on the collocational knowledge of CFL learners can be drawn from these results.

Acknowledgments

I would like to express my sincere gratitude to my supervisor Prof. Susanne Weigelin-Schwiedrzik for her suggestions, advice and guidance, which have helped me immensely while working on this thesis.

My great appreciation also goes to Chen Tingting, Ph.D. who allowed me to replicate her study and assisted me when I had questions on details of her study.

I would like to acknowledge the professors at the Chinese department who allowed me to collect data during their course hours, Prof. Susanne Weigelin-Schwiedrzik and Mag. Patrice Gruber, Bakk., without whom data collection for this study would not have been possible.

I would also like to thank my best friends, Petra and Gözde, who have always supported me in everything I do.

Finally, I would like to thank my family for always providing me with unfailing support and encouragement throughout my life and my studies. This accomplishment would not have been possible without your love and support.

Table of Contents

Abstract	iii
Acknowledgments	v
1. Introduction	3
2. The lexicon in general linguistics	4
2.1. What is a 'word'?	4
2.2. Multiword units	9
2.3. What is a 'collocation'?	12
2.4. Organisation of the mental lexicon	14
2.5. Frequency aspects of the lexicon	18
3. The Chinese lexicon	21
3.1. What is a 'word' in Chinese?	21
3.2. Word classes	24
3.2.1. Content words	25
3.2.2. Function words	29
3.3. Collocations in the Chinese lexicon	30
3.4. The Chinese writing system	36
4. Learning vocabulary	39
4.1. What does it mean to 'know a word'?	40
4.2. Dimensions of word knowledge	43
4.2.1. Breadth of vocabulary knowledge	43
4.2.2. Depth of vocabulary knowledge	45
4.3. Types of vocabulary learning	49
4.4. Important factors for learning vocabulary	51
4.5. Learning vocabulary in CFL	55
5. Assessing vocabulary knowledge	61
5.1. Assessing vocabulary breadth	66
5.1.1. Yes/No Test	66
5.1.2. Vocabulary Levels Test	67
5.1.3. Vocabulary Size Test	68
5.2. Assessing vocabulary depth	70
5.2.1. Vocabulary Knowledge Scale	71
5.2.2. Word Associates Test	72
5.2.3. V_Links	74
6. Research methods	76
6.1. Model study	77
6.2. Adapted test	81
6.3. Context of the study	84
6.4. Data collection and analysis	86
7. Results	88
7.1. Demographic questionnaire	88
7.2. Word Associates Test (WAT)	91
7.2.1. General identification of collocations	93

7.2.2. Identification of collocations in relation to word frequency	98
8. Discussion and Conclusion.....	112
8.1. Discussion of the results	112
8.2. Implications for teaching and learning CFL.....	117
8.3. Limitations.....	119
8.4. Conclusion	120
Bibliography	122
Appendix A Chinese Word Associates Test (WAT).....	128
Appendix B Demographic background questionnaire.....	137
Appendix C List of tables, illustrations and figures	140
Appendix D Abstract (German)	142

1. Introduction

Across the globe it is an advantage to know and be able to use a foreign language. Of course certain languages are more wide spread than others, however, the importance of certain languages can change due to the political and economic status of their origin country. For decades, English has been and still is the predominant foreign language worldwide. Nevertheless, Mandarin Chinese has become more and more important due to the economic rise of China after the economic reform, which started in 1976. As a consequence Mandarin Chinese has become a sought after foreign language skill by many businesses. Unsurprisingly, the number of people learning Chinese as a foreign language (CFL for short) is increasing steadily.

Learning a (foreign) language however encompasses many different aspects. A crucial part is learning a language's lexicon or vocabulary, as a certain number of words is necessary to be able to communicate in a language. However, the size of vocabulary is not the only relevant factor here because what information we know about certain words is also crucial. In the field of CFL neither of these aspects is as well researched as in the fields of other major foreign languages (especially English as a foreign language). This is the reason why this MA thesis focuses on the area of vocabulary depth of CFL learners. More specifically, the aim of this study is to examine the vocabulary depth of MA students of Chinese studies at the University of Vienna. As vocabulary depth per se is still a rather huge field to examine, this study turns its focus on collocations in the Chinese language and how well the participants can identify them. It should be noted here, that this study is replicating the study by Chen (2016), which she conducted for her PhD thesis.

Chapter two provides insight into the relevant literature on the (mental) lexicon in general linguistics, considering questions such as what a word and a collocation is or how the mental lexicon is organised. Chapter three on the other hand considers the questions of what a word and a collocation is in Chinese and also gives information about other aspects of the Chinese language that are relevant to this study (e.g. the Chinese writing system). Learning vocabulary is discussed in chapter four, which also includes different dimensions of word knowledge and different types of vocabulary learning. Consequently, chapter five discusses the assessment of vocabulary knowledge. This includes different types of vocabulary tests – for both vocabulary size and depth. However, it should be noted here that the 汉语水平考试 [Hànyǔ shuǐpíng kǎoshì] (HSK), will not be discussed as this is a language proficiency test for the Chinese

language and is, thus, not a test that was specifically developed to test said knowledge. Chapter six will then discuss the research methods used to conduct this study, while chapter seven will discuss its results.

2. The lexicon in general linguistics

The topical area forming the foundation of vocabulary and vocabulary knowledge is: the lexicon. The term *lexicon* describes the set of words of a language, which is abstract and stored in our mind. A concrete representation or description of said lexicon can be found in the form of a dictionary. And although these two entities have a relationship, their actual structures do not correspond and there is little similarity between them. A dictionary's organisation is heavily influenced by practical considerations, such as ease of consultation, which is why the words in a dictionary are listed in alphabetical order. The lexicon, however, is not stored in this alphabetical, word-by-word fashion; more on the storage of the lexicon will follow later on in this chapter (in section 2.4. *Organisation of the mental lexicon*). Another term that is related to this area as well is the term *vocabulary*, which can be ambivalent in this context because it can be used to denote either the body of words in use in a particular language (i.e. the lexicon) or the collection and description of these words (i.e. a dictionary). (Aitchison, 2012; Ježek, 2016)

The lexicon is the focus of different disciplines, which study different purposes and parts of the lexicon. Lexicology, for example, studies properties of words, as well as their relation to each other. One of the goals of lexicology is the development of a theory of the lexicon. A more practically oriented discipline studying the lexicon is lexicography, as its focus is the compilation of dictionaries. (Ježek, 2016)

2.1. What is a 'word'?

For non-experts the definition of the term *word* might be relatively easy and intuitive but for linguists it is quite a difficult challenge. A starting point to ease into this complicated area is looking at the most primary characteristic of a word, which is its information content or meaning. But what counts as a word meaning, how can word meanings be defined and represented? This is a question, which scholars of many disciplines, such as linguistics or philosophy, have dealt with. We will now focus on how meaning can be associated with words. In a language two dimensions can be identified: form and content, with it being possible to isolate single elements:

the lexical form <earth> [3:θ] and meanings, such as “planet in the solar system”, for example. But there is also the case of meanings that have not been associated with a single lexical form. These can be thought of as *concepts*, which are mental categories carrying some information content but whose existence is independent from language. The process of association of a concept with a lexical form, which results in the existence of a word, is called *lexicalisation*. So following this line of thought every word in a language is the outcome of lexicalisation. But this is a quite broad interpretation of this term. According to a narrower interpretation lexicalisation “refers to the specific process according to which a sequence of words that frequently recur together in texts acquires the status of lexical unit with an autonomous meaning” (Ježek, 2016:6). This is a simplification of the lexicalisation process, as this is quite complex, with different lexicalisation patterns existing and rarely a one-to-one correspondence of a word and a concept. But meaning is not only a feature of words. Other elements of language can also bear meaning: syntactic structures, such as ‘the passive’, or morphological categories, such as gender or number, also carry meaning. What is the difference between the meaning of a word and that of syntactic or morphological units? One clear difference is that the speaker can more easily perceive and describe the meaning of a word than that of syntactic or morphological units. To illustrate this point, the distinction between content and function words shall be considered. This distinction is generally aided by including word classes and showing, which word classes belong to which type of word. *Content words* (also called *lexical words*) include nouns, verbs, and adjectives, while *function words* (also called *grammatical words*) include determiners, pronouns, conjunctions and prepositions. Adverbs can belong to either category: content adverbs, such as ‘quickly’ belong to content words, while others, such as *not*, belong to the function word category. The meaning of content words, which is called *lexical meaning*, is (to a certain degree) semantically autonomous. The meaning of function words on the other hand depends on the content words they are related to. Thus, this meaning is called *grammatical meaning*. Consider the English word ‘in’ as an example: in the phrase ‘in the morning’ it means ‘during’ because ‘morning’ denotes a time period, while in the phrase ‘in the building’ it means ‘inside’, as ‘building’ denotes a place. Content and function words can be further distinguished by the following characteristic. The former can be seen as an open set of items that gains and loses members relatively quickly. Because a person can easily invent new words, the acquisition of new content words into the lexicon is a recurring phenomenon. The latter, however, is a closed set, which tends to remain constant through time, as the occurrence of new function words is a more isolated, less frequent phenomenon. If a new function word does occur, it is often by transforming an already existing lexical word into a function word, for example. Ježek (2016:15) takes the English word ‘while’ as an example: originally it denoted a

noun with the meaning 'moment' or 'period of time' (e.g. 'in a little while'), later on it took on the meaning of a conjunction with the meaning 'during the time that' (e.g. 'while she was writing, ...'). (Aitchison, 2016; Ježek, 2016; Plag, 2003)

However, grammatical meaning is not only carried by function words. As mentioned above, other elements can also express grammatical meaning (e.g. syntactic structures and grammatical morphemes). One syntactic structure expressing grammatical meaning is the passive. An example in English is the sentence 'the book was stolen'. By using the passive, the focus is removed from the entity that brings the event about (in our example this is the person who stole the book). This syntactic structure's meaning is to present a situation from the point of view of someone who is not actively participating. The following are additional types of grammatical meaning: number, gender, time, aspect and voice. (Ježek, 2016)

Having discussed how words 'come into being', the counting of words in a language is the next aspect we should focus on, as this is not an easy task either. This difficulty is better illustrated by looking at an example: the English word 'walk'. Besides the word 'walk' there is a whole family of forms that go with it, such as 'walks', 'walked', 'will walk' and 'has been walking'. All these variants are inflected forms because English is a (weakly) inflectional language. So the question now is how should the inflected forms be analysed? The answer is that they are different forms of the same word and can be ascribed to a single word. Now a new and more difficult question arises: which form represents the word? This form, which is called *citation form*, is determined by convention. For the English language, the convention is as follows: the infinitive for verbs, the singular for nouns or adjectives. But it is important to remember that we could also choose a different form to represent a word, which may be the case in other languages. (Aitchison, 2016; Ježek, 2016)

Two concepts are important when following this line of reasoning and trying to account for the existence of form families for single words. These are the *lexeme* and the *lemma*. A lexeme corresponds with the base form to which the inflected forms are ascribed, as seen in the example above. A lemma on the other hand is a lexeme seen as a dictionary entry, which is why it is also called *lexical entry*. Furthermore, a lexeme does not only include a form family but also category families (such as nouns, verbs, etc.). Consider the English word 'fish', for example: 'fish' as a noun denotes an animal living in water but 'fish' as a verb denotes the action of catching said animal. The discipline of lexicography deals with this in the following way: 'fish' (noun) and 'fish' (verb) belong to the same headword, namely fish' but they are two lemmas of this headword. So here the term lemma contains all the inflected forms of a

headword belonging to the same syntactic category. A further question arising in this line of reasoning is how derived words are categorised. Derived words (e.g. 'art' – 'artist') share the same lexical morpheme (in this example 'art') but they are two distinct words because while having a different form they also carry a different lexical meaning. The same is true for compound words, such as the English words 'sunglasses', 'bookstore' or 'bedroom'. They too are distinct from their base words (i.e. 'sun', 'glass'; 'book', 'store'; 'bed', 'room' in this example). Another category that has to be taken into account is the one consisting of *multi-word lexical units*, also referred to as *multi-word expressions* or *lexicalised phrases*. Two different types of multi-word lexical units exist: *idiomatic expressions* and *idiomatic constructions*. These two types of expressions differ in the following way: the term idiomatic expression refers to expressions whose meaning is non-compositional (such as 'round table' for example), whereas the term idiomatic construction refers to expressions exhibiting internal variability (i.e. one or more open slots may be filled by different words, producing several variations of the same expression, as in 'bring an issue / a concern / a fact into light'). Considering the points covered above, it is evident that counting the number of words in a language is not as easy as one might think because quite a number of factors need to be taken into account (inflected or converted forms, derived words, etc.). (Ježek, 2016)

Another perspective on the term *word* beyond attempting to establish their number is by looking at their internal structure, which we will call *word types*. Words with an internal structure generally fall into two main groups: group one consists of words, which have been formed by applying the available word-formation rules of a language (i.e. rules of derivation, compounding, conversion) and group two is made up of words, which come from syntactic and semantic links existing between two or more words, which occur together in text.

First we will focus on the words of group one. Based on their form, words can be classified into two distinct types: *simple* and *complex words*. Simple words, such as the English words 'table' or 'happy', consist of a single morpheme. In contrast, complex words consist of more than one morpheme (besides inflectional endings) and therefore have a morphological or syntactic internal structure. Derived words and compounds are examples for complex words with a morphological internal structure. Derived words are formed by a lexical morpheme and by adding one or more affixes, as is the case with the word 'arrival' for example (which is composed of 'arrive' and '-al'). Compound words are formed by a minimum of two independent morphemes with lexical meaning. 'Bookstore' and 'sunglasses' would be two examples for

compound words. Complex words with a syntactic internal structure, on the other hand, look like phrases but exhibit internal cohesiveness that ordinary phrases lack.

In order to identify whether a sequence of words works the same way a simple word several empirical tests exist. With the help of these tests the following properties can be verified that are essential to a simple word. Said properties are cohesiveness and fixed order of the constituent parts, which are based on the assumption mentioned above that words have an internal structure (i.e. a word can be broken down into phonemes, syllables and morphemes). However, they still “constitute single units at some level of linguistic analysis” (Ježek, 2016:25). Consider the *separation test* as one example of these tests for “wordhood”. To perform this test the parts that are supposed to be the constituents of a word are separated by putting lexical material between them (e.g. an adjective to modify a noun if there is a noun). The result is then evaluated regarding its acceptability: if the result is not acceptable, the parts are said to constitute a word because words cannot be partially modified or separated. In the following a few examples are discussed for better illustration. The English compound ‘waiting room’ has a word-like status according to the separation test as a modification, such as ‘waiting little room’ is not acceptable. The same goes for the word ‘credit card’: ‘a credit new card’ is not acceptable. These two expressions can only be modified globally, i.e. ‘a little waiting room’ or ‘a new credit card’. Expressions such as ‘take care of’ or ‘take advantage of’ have different results, as it is acceptable to say ‘take good care of’ or ‘take full advantage of’. A second test of “wordhood” is the *substitution test*, which can be used by substituting one of the constituents of a sequence of words for a (near-) synonym. Consider ‘blind spot’ as an example again. It is not acceptable to substitute ‘spot’ with a near-synonym, such as ‘place’ or ‘location’ (i.e. ‘blind place’ and ‘blind location’ are not acceptable). The expression ‘blind spot’ has word status, as a substitution is not allowed by ordinary phrases. (Ježek, 2016)

Group two consists of words, which develop through syntactic and semantic links between two or more words. Two main types of word combinations exist: *free* and *restricted combinations*. Free combinations are, as their name suggests, word combinations without restrictions applying to them. However, no combination exists that is completely free of restrictions. Restricted combinations on the other hand can be divided again into two main types according to linguistic constraints. Combinations can be restricted based on their lexical solidarity or based on their patterns of use. The latter will be the focus of the next section (i.e. 2.2. *What is a ‘collocation’?*). Therefore only the former will be discussed in the following. Four main observations can be made about combinations that are based on their lexical solidarity. First of all, these

combinations show varying levels of generality or specificity regarding the taxonomy of categories. This means that the target of these combinations may be a general category, such as a *concrete object*, or a specific category, such as *vehicle*. Consider noun-adjective combinations as an example: some adjectives can be applied to several noun classes (e.g. 'long': 'long hair', 'long day', etc.), some can be applied to restricted sets of nouns (e.g. 'blonde': 'blonde hair' or 'blonde beer'), some can be applied to a single class (e.g. 'isosceles': 'isosceles triangle') or to a single object (e.g. 'curdled': 'curdled milk'). Secondly, their meaning is compositional, which is why their meaning can be predicted from the meaning of the single words which were combined. The third observation is that the possibility of substitutions for the members of the combination is constrained by at least one restriction. When this is a specific restriction, a substitution of a member becomes impossible (e.g. 'curdled milk': only milk can be curdled). Lastly, from a syntactic point of view, the 'members' of the combination behave like free elements because they may undergo various syntactic operations, such as passivisation (e.g. 'the book can be bought') or relativisation (e.g. 'the book that I bought'). (Ježek, 2016)

2.2. Multiword units

To provide a single all-inclusive definition, classification and terminology of *multiword units* is difficult, as they come in so many different forms and their nature is rather complex. Nation (2013:479) lists four kinds of multiword units: (1) a group of words that commonly occur together (e.g. *take a chance*), (2) a group of words whose meaning is not directly derivable from the meaning of its parts (e.g. *be taken in*), (3) all combinations of a particular (type of) word and its accompanying words (may they be of high frequency, strong association or not) and (4) a group of words seen as intuitive formulaic sequences. A variety of criteria are the basis of different kinds of multiword units, which include frequency of co-occurrence, compositionality, form and storage. This is also the reason why the list of terms that cover these units is constantly growing. In Wray's study (2000) 50 different terms can be found, which include *composites*, *idiomatic phrases*, *phrasal expressions* and *routine formulae*. These various terms indicate the different ways researchers look at multiword units and that the research in this area is not uniform but fragmented. Consequently it is necessary to set clear criteria and then develop standard terminology for the different kinds of multiword units, at least where it is possible. But the definition of what multiword units are also has a practical impact on what appears in learners' dictionaries. (Nation, 2013; Walker, 2011)

If it is already this difficult to come to a definition of multiword units, can we find characteristics for them? It is indeed the case that most, if not all, multiword units share certain characteristics.

First of all it can be said that the nature of most multiword units is variable, which means they can change their word order, their occurring words or their morphological form of their collocates. Consider the collocation of 'save' and 'skin'. According to Sinclair (2004a:275) the Bank of English shows 201 examples of (any form of) the verb 'save', which is followed by 'skin' within three words. In the instance where only one word intervenes between 'save' and 'skin', it is a possessive most of the time (with 'her' and 'your' being especially prominent). If more than one word intervenes, the majority of the time it is a combination of a possessive and the word 'own' (e.g. 'save her own skin'). Additionally, it is also possible to replace 'skin' by another word (e.g. 'hide', or in American English 'ass' or 'butt'). Secondly, the words that make up such a unit are not combined at random, but in unison with typical grammatical and semantic uses of these words outside the unit. Additionally, for some multiword units it is true that their meaning goes beyond what their individual parts mean. These units are made up of multiple words and literature suggests that multiword units are probably stored as single choices. And lastly, multiword units are used for communicative purposes in the same way as single words. (Nation, 2013; Sinclair, 2004a; Sinclair, 2004b)

The reasons for the importance of multiword units in language use are strongly related to the nature of language, language knowledge and language use. Language generally speaking is a sequence of language items, which is why a lot of language learning and language processing is related to associations between these sequential language items. A major factor affecting this way of learning is the frequency of encounter. Two models of how words occur in a text: the *open-choice principle* and the *idiom principle*. The open-choice principle sees language text as a series of choices with the only limitation being that the choices have to be grammatical. With the idiom principle on the other hand there are more limitations and constraints because the choice of a word has a limiting effect on the following word choices. These limitations can come from the nature of the world, choice of register and the collocational knowledge of language. Multiword units are an important goal of learning a language because a large amount of any language follows the idiom principle. (Nation, 2013; Sinclair, 2004a)

We have seen that multiword units play an important role in language but if we want to study them systematically we need to identify them. Using criteria based on form, meaning and storage, three major approaches to identify these units exist. First we will discuss *form-based approaches*, which use frequency of co-occurrence of the words making up a multiword unit. The methods these approaches mainly use are corpus-based searches to find words that frequently co-occur. For language learners the frequency of occurrence is an important factor

because they need to learn items that they will read or hear often and therefore also need to use often. But frequency of co-occurrence cannot be the only criterion used to identify multiword units because then units such as *of the*, *in the* or *one of the* would be at the top of a multiword unit list. Consequently there are other criteria that need to be applied in order to identify 'real' multiword units, which we mentioned shortly above. The following factors affect counting the frequency of multiword units. *Adjacency*, which describes that words of a multiword unit can occur right next to each other, but also be separated by one or more words that do not belong to the unit. This fact makes it important to be careful and check which items belong to the unit and which do not (e.g. 'by and large' or 'serve sb. right'). *Grammatical variation* is another factor – that is the form of a multiword unit can be fixed or variable (e.g. 'to and fro', 'pulling your leg'). If variable items are included, a variety of search words and combinations of words need to be included. *Grammatical well-formedness or structure* is important as well because such a unit can be a complete grammatical unit (e.g. a sentence, a sentence subject or predicate) but it can also be grammatically incomplete (e.g. 'on the other hand', 'on the basis of'). Multiword units can be also *lexically (in-)variable*. Some words of multiword units can be substituted by words of related or similar meaning (e.g. 'once a week', 'twice a week', 'once a month'), others cannot (e.g. 'as well as'). (Nation, 2013)

Secondly, *meaning-based approaches* consider the meaning of the parts of a unit and how transparently they make up the meaning of the whole unit. Multiword units can be rated on a scale by using the criteria of compositionality and figurativeness. This means whether the parts make up a whole and whether an item should be rated figuratively or literally. This rating scale includes three points: *core idioms* where the meaning of the whole cannot be interpreted from the meaning of its parts (e.g. 'kick the bucket', '(raining) cats and dogs', 'pull the other one'), with *figuratives* items need to be interpreted figuratively (e.g. 'out of the blue', 'a bone of contention', 'a stumbling block') and with *literals* the meaning of the whole can largely be interpreted by the meaning of the parts (e.g. 'at the moment', 'you know', 'take a vacation'). (Nation, 2013)

Thirdly, *storage-based approaches* consider the storage of multiword units in the brain. Some questions asked by these approaches are: Are (some) multiword units stored in the brain as whole units? Which multiword units are stored and why? Are language use and language learning affected by such storage? These approaches are different from the two discussed previously, as corpus data and classification by meaning are insufficient criteria in this case because storage in the brain is a psychological phenomenon. (Nation, 2013)

2.3. What is a ‘collocation’?

This thesis tests the ability of the study’s participants to identify collocations in Chinese (more on the study will follow in chapter six *Research methods*). A discussion of the notion of ‘*collocation*’ is therefore central to the present chapter. More on collocations in the Chinese language will follow in a section of the next chapter (3.5. *Collocations in Chinese*).

This section connects directly to knowing a word (i.e. 4.1. *What does it mean to know a word?*) section and to multiword units (i.e. section 2.2.) because knowing a word includes knowing which words it typically co-occurs with and because a collocation is a particular type of multiword unit. Collocations can differ in size, type, closeness of collocates, as well as in the variability of collocates, aspects which we already discussed in the section on multiword units.

Not unexpectedly it is quite difficult to define the term collocation because several different definitions exist. We will start by looking at broad and narrow definitions of the term *collocation*. A very broad definition states that a collocation is any frequent co-occurrence of two words in a language. Sinclair’s (1991:170) definition is in line with this broad definition, as he defines collocations as “the occurrence of two or more words within a short space of each other in text” that are “frequently repeated”. In the broad definition the co-occurrence of collocates reaches a statistical significance but it ignores the grammatical and semantic limitations of collocations. Also, according to Benson, Benson and Ilson (1986) collocations are arbitrary, recurring combinations of words, making arbitrariness and repetition two important factors for collocations. However, according to Sun, Huang and Fang (1997:30) and Zhen, Ren and Yin (2006:82) collocations also follow a certain structure and are domain-dependent. Sun, Huang and Fang (1997) define as collocations combinations that are not part of common and daily communication but are part of special field, such as a certain area’s special technical terminology and idioms.

Generally it can be stated that word co-occurrence is limited by at least one or more restrictions, as lexical and grammatical factors need to be considered as well. This definition is more precise than the first one, as it shows that the co-occurrence is not merely statistical but based on a restriction, which favours this word combination and at the same time restricts their compatibility with other words. Thus, according to Ježek (2016:200) a collocation is a word combination subject to restrictions but opposed to the above definitions, this linguistic view also includes the fact that the choice of a particular word, which is called *collocate*, to express a given meaning is

influenced by the word it applies to, which is the *base* of the combination. The English combination *heavy rain* can serve as an example. The noun *rain* is the base of the collocation, while the collocate *heavy* refers to the intensity of the base.

Now the question arises what is the difference between a lexical semantic solidarity (e.g. *to wear* and *garment*) and a proper collocation (e.g. *pay attention*)? While they indeed are similar phenomena, some distinctions between them can be drawn, which can be seen if the following test is applied: can a foreign speaker of a language guess the noun from the verb or adjective, if he or she only knows the meaning of words but not of the collocations? With lexical semantic solidarities one of the terms, even when it stands on its own, implies a certain second term. For example, even standing alone the verb *to wear* implies the noun *garment* or the adjective *blonde* implies the noun *hair*. With proper collocations, however, this is not the case, as the connection is only present in the combination but is absent if the terms of the collocation are looked at individually. The verb *to pay* does not necessarily imply the noun *attention* or the noun *bill*, for example. The reason for this is that the collocate can describe a property or action that may apply to several classes of objects or that it no inherently expresses an activity in which the referent of the base participates. For example, *heavy* does not necessarily refer to the intensity of all base words it can be combined with, as it does in the collocation *heavy rain*. The solidarity with the base is only established in the specific use. The direction of restrictions differs between lexical semantic solidarities and collocations as well. In the former the verb and adjective impose certain semantic requirements on the noun. In collocations, however, the restriction moves in the opposite direction, which means that the noun is the one calling for a specific verb or adjective. (Ježek, 2016)

When looking at the syntactic context, different types of lexical collocations can be distinguished. According to Benson, Benson and Ilson (2009:XXXI-XXXIV) a total of seven major types are distinguishable. The first type is the collocation of a *verb denoting creation or activation and a noun, pronoun or prepositional phrase*. Collocations with verbs of creation are 'to make an impression', 'to set a record' or 'to inflict a wound', while 'to set an alarm', 'to fly a kite' or 'to launch a missile' are collocations with verbs of activation. In some instances, the same noun can be used with both types of verbs (e.g. 'to apply a principle' –activation and 'to establish a principle' –creation). In a lot of cases the acts of creation and activation are united in one verb (e.g. 'to impose an embargo' or 'to display bravery'). Because this type of collocation can be rather arbitrary and non-predictable, they are hard to deal with for the learners of a language. The second type of lexical collocation is the co-occurrence of a *verb denoting*

eradication or nullification and a noun, such as ‘to reject an appeal’, ‘to lift a blockade’ or ‘to demolish a house’, for example. The co-occurrence of *adjectives and nouns*, such as ‘strong tea’ or ‘weak tea’, is the third type of lexical collocation distinguished by Benson, Benson and Ilson (2009). It is important to note that more than one adjective or more than one form of the same adjective can form a collocation with the same noun: the adjectives ‘warm’, ‘warmest’, ‘kind’, ‘kindest’ and ‘best’ can all collocate with the noun ‘regards’. In English nouns can also be used attributively, such collocations fall into this category as well, if the second noun keeps its basic meaning (i.e. the meaning it has on its own). ‘Land reform’ and ‘house arrest’ are example of such collocations. The fourth type is collocations of *nouns and verbs in which the verb denotes the action characteristic of the entity* (e.g. person, thing) *designated by the noun* (e.g. ‘adjectives modify’, ‘alarms go off’ or ‘blood circulates’). The next and fifth type of collocation denotes *a unit associated with the noun* in the form of noun₁ of noun₂. This can either be a larger unit to which a single member belongs (e.g. ‘herd of sheep’ or ‘bouquet of flowers’) or a specific, concrete, smaller unit of something larger (e.g. ‘a piece of advice’ or ‘an article of clothing’). The co-occurrence of *adverbs and adjectives* is the sixth type of lexical collocation, such as ‘deeply absorbed’ or ‘closely acquainted’. The seventh and last type of lexical collocation is the combination of *verb and adverb*, such as ‘affect deeply’, ‘appreciate sincerely’ or ‘argue heatedly’.

Concerning collocations the following factor is also important: *salience*. If we consider the example *rain* again, there are two expressions: *heavy rain* and *pouring rain*. According to Ježek (2016:203) *heavy rain* appears about four times more frequently than *pouring rain* but the score for saliency is higher with *pouring rain*. Although *heavy* and *rain* appear together more often, *heavy* occurs with a large number of words, while *pouring* can be combined with only a restricted set of words.

2.4. Organisation of the mental lexicon

In the introduction to this chapter, the differences between the organisation of the mental lexicon and a dictionary were briefly mentioned. After discussing what words, multi-word units and collocations are, we will now turn our attention back to the organisation of the mental lexicon. While common dictionaries we know are organised alphabetically, our mental lexicon does not tend to work that way, although certain aspects of word sound structure, such as initial sound, ending, stress pattern or stressed vowel, are likely to play a role in word storage. Evidence can be drawn from slips of the tongue. Slips of the tongue occur when people make mistakes selecting one word instead of another. If the mental lexicon was indeed organised in

alphabetical order, it would be expected that an adjacent entry of the actual word was used. Then we would expect the words *zit* (i.e. a spot on the skin) or *ziti* (i.e. pasta in the form of tubes) instead of the musical instrument *zither* because the two terms precede and follow it in a dictionary. But such mistakes are highly unlikely. Looking at Aitchison's (2012:12) examples of slips of the tongue, such as "*he told a funny antidote*" (with *antidote* instead of *anecdote*), or "*the doctor listened to her chest with his periscope*" (with *periscope* instead of *stethoscope*), we can see that some aspects of sound structure indeed play a role in the organisation of the mental lexicon with the order of the words not being straightforwardly alphabetical. However, meaning is taken into consideration as well because words with similar meanings are confused often, as the following example shows: the speaker wants to crack a nut and says, "*Please hand me the tin-opener*" when really the speaker must have meant "*nut-crackers*" (Aitchison, 2012:12). While the content of dictionaries is fixed, containing a certain number of words, the content of the lexicon is not fixed. This means that new words can be added to it all the time, pronunciations or meanings of already existing words can be altered and we can also think of new words. This shows that the mental lexicon is fluid and flexible. Comparing the mental lexicon and dictionaries further, it becomes noticeable that in dictionaries words mostly appear in an isolated context because most of the time dictionaries only include the definition of a word but no information on how this word relates to other words or on how it is connected with other words. For example, a dictionary contains the words *child*, *infant*, *baby* and *toddler*, with their respective definitions but will not show the relations and connections between them. Dictionaries also contain little data regarding the syntactic patterns of a word. All this information, however, is an intrinsic part of the mental lexicon, which shows that it contains a lot more information than any dictionary. We have mentioned above that dictionaries list words in alphabetical order for practicability reasons. Researchers agree that in the mental lexicon words are probably stored in word groupings. The members of these word groupings share similarities, such as form or meaning. Typical lexical structure would be storing words based on: morphological word families, such as *book*, *booking*, *booklet*, and *bookstore* (based on *book*); semantic networks, for example *buy*, *sell*, *negotiate*, and *pay* (based on meaning associations); or based on similar syntactic behaviour, such as nouns, adjective or verbs. (Aitchison, 2012; Ježek, 2016)

Because the mental lexicon is stored in the brain and therefore not directly accessible, psycholinguists have had to use other tools to make a model of it. Various types of these clues to the mental lexicon exist: word searches and slips of the tongue, speech disorders, brains scans, linguistics and linguistic corpora and finally psycholinguistic experiments. Consider the

first clue of word searches and slips of the tongue. The stages of word searches give us clues about the organisation of a whole area, as “words which seem to be closely related may be stored close together, or they may be more distant but have strong links connecting them.” (Aitchison, 2012:20) Sometimes when searching for a word, we might be completely blocked and unable to remember anything about the word we are looking for, while in most cases we remember something about it that will eventually lead us to the word we are looking for. Aitchison (2012:20) shows us an example for rooting around for a certain word: “He took a ... er ... something interest in it. What’s the word I want? It’s something like *salient*, but I may be confusing it with *prurient*. I know it begins with an s. Ah, yes, SALACIOUS [sic].” Despite providing evidence for the organisation of the mental lexicon, word searches also have drawbacks. Most mislaid words tend to be relatively uncommon, which is why the method of searching might be different for frequently used words. A second drawback would be that the straightforward way of searching for a word was blocked and therefore the speakers took an unnatural route to the target word they were looking for. The evidence we get from speech errors, such as slips of the tongue, is valuable because on the one hand it is unlikely that speakers approach their target word by a roundabout route because they think they have chosen the right word, at least momentarily. On the other hand, slips of the tongue happen to everybody no matter their level of education, which is why this reflects the working of the normal brain. Additionally, slips of the tongue follow certain patterns. There are two major types of slips of the tongue: *assemblage errors* and *selection errors*. Examples for the former are errors such as *patter-killer* (instead of *caterpillar*) or *par cark* (instead of *car park*). This shows that the right items were chosen by the speaker but assembled in the wrong order. In the latter, wrong items are selected from the mental lexicon, such as *corporal punishment* (instead of *capital punishment*). These selection errors can be based on different similarities between words: meaning similarities, sound similarities or even both. However, not only single words can replace the target word but *blends* can also occur. These are cases where two words are blended (i.e. combined) into one word. Examples for both, single errors and blends, can be found in the table below:

	SINGLE ERRORS	BLENDS
MEANING	He came <i>tomorrow</i> (yesterday).	I don’t <i>expose</i> (expect + suppose) anyone will eat that.
SOUND	The emperor had several <i>porcupines</i> (concubines).	He was a <i>lustrious</i> (lustful + illustrious) and

passionate man.

MEANING	You can hear the	
+ SOUND	<i>clarinets</i> (castanets) clicking.	My <i>tummach</i> (tummy + stomach) feels funny.

Table 1. Types of selection errors (Aitchison, 2012:22)

As mentioned above, further evidence for a model of the mental lexicon can come from speech disorders. Aitchinson (2012:24-26) lists severe speech difficulties caused by a head injury or a stroke, called aphasia, as one example. The most common symptom of most types of aphasia is word-finding difficulties. The following are common examples of mistakes patients with aphasia make: *chair* (for *table*), *knee* (for *elbow*) or *hair* (for *comb*). The reason why aphasia is interesting for researchers is that they assume that certain patterns of symptoms are only possible if the cognitive system is organised in a certain way. Patients whose mental lexicon is selectively impaired (i.e. some parts are damaged, some are not), for instance, can give indications for subsystems within the mental lexicon. For example, if an aphasic could remember nouns but not verbs, this would indicate that they might be organised differently in the mental lexicon.

Another source for clues about the organisation of the mental lexicon are psycholinguistic experiments. Various types of tests or experiments are used to gather information about how we store language in our brain. One popular method are word association experiments: the test subject is presented with a series of words and has to name the first word coming into mind for each item. For example, the researcher says: "What do you think of first when I say *day*?" The subject might answer *night* or *light* or something else. This type of experiment is useful because different people still tend to give rather similar responses. Another popular type of experiment for investigating the mental lexicon is the lexical decision task. The subjects are presented with a number of sound and letter sequences and are asked to decide whether they are words or not. Reaction times are measured (in milliseconds) because this may provide information on words most readily available in their mental lexicon. However, the distinction between word-finding time and various other processes that might be happening in the mind (after a word was found but before a response has been given) needs to be considered here. Different variations of the lexical decision task exist: reaction time to common words compared to reaction time to uncommon words; rejection speed for nonsense words (comparing nonsense words that are similar to real words and nonsense words that are not); or looking at how words presented before it, may alter the response to a word. (Aitchison, 2012)

Although the organisation of the mental lexicon in general is very intriguing and much more can be said about it, at this point we should also consider the mental lexicon of one person when more than one language is stored in it, which is true in most cases nowadays. A question here certainly is whether L1 and L2 share the same lexical store or whether they are stored differently. The way of lexical storage of L1 and L2 is directly related to language proficiency. Studies provide us with evidence, as they show that at low proficiency levels L2 words and their L1 equivalents are directly connected. This happens regardless of the way the words were learned. However, new words can also be accessed without going via the L1, by accessing the meaning of the word directly. Experiments also show that between-language connections are made in different ways: for aspects of meaning, which are part of the core concept of a word and therefore are independent from context, connections between languages are made. For context dependent aspects of meaning (e.g. collocates), however, this is not the case. These aspects of meaning are not transferred from the learner's L1 but can only be learned through experience and application of the L2. (Ježek, 2016; Nation, 2013)

2.5. Frequency aspects of the lexicon

The definitions of the notions *word*, *multi-word units* and *collocations* have been discussed in this chapter. However, the aspect of *word frequency* in the mental lexicon should not be ignored. Word frequency is based on statistical properties of language, which is why the question arises whether humans are able to intuitively determine the underlying frequencies or tendencies of language. According to literature (e.g. Divjak, 2012) humans are indeed able to determine such underlying frequencies: statistical properties of linguistic input can be used by learners to discover structure (including sound patterns, words and the beginnings of grammar). This leads to the hypothesis that not environmental triggers but a probabilistic, language experience based grammar is used by humans to specify a set of linguistic properties. The activation or probability values of one's cognitive organisation of language, as well as linguistic categories and linguistic structures are determined by their relative frequencies in language use. The reason why humans have statistical information about word combinations' frequencies is that they store exemplars, which aids acquisition, processing and inferring grammatical structure. (Caldwell-Harris, Berant & Edelman, 2012)

The disciplines dealing with the effects that frequencies of language use have on cognition can be said to be an intersection of cognitive corpus linguistics, psycholinguistics and computational modelling. While corpus linguistics is interested in frequency variation across utterances, psycholinguistics is more interested in theoretical issues and usage-based models.

Computational modelling tries to implement frequency effects together with grammatical and lexical regularities into their models. Many studies on the theoretical impact of frequency of a psycholinguistic nature do not include (corpus) linguistics. The readily identifiable forms are considered without their wider linguistic context. On the other hand, corpus linguistic studies have tried to map speakers' minds with textual datasets, ignoring the fact that learning by probabilistic means is not strictly proportional to the stimulus, as frequency is not the only factor. So far, identifying frequency effects in a number of areas has been the focus of research on frequency. The existence of such frequency effects have been documented for many areas of linguistic analysis (i.e. phonology, morphology, syntax, lexicon, etc.). Research has also shown that frequency plays a facilitatory role in building the basic structures of a language. (Divjak, 2012; Caldwell-Harris, Berant & Edelman, 2012)

Researchers have distinguished two approaches to the mental lexicon in connection with grammar, which also have different views on the role frequency plays in the mental lexicon. On the one hand *dual-system theories* exist (e.g. Chomsky, 1965, 1995). On the other hand there are also *single-system theories* (e.g. Bates and MacWhinney, 1989) to be found. The former see a clear distinction between the mental lexicon (i.e. an inventory of memorized forms and mental grammar) and grammar (i.e. rules and constraints on combining these memorised items). These theories assume that the mental lexicon contains linguistic units, such as simple words (e.g. 'cat'), morphemes, irregular nouns and verbs, as well as longer non-compositional phrases (e.g. idioms) (i.e. non-derivable linguistic units). One model of the dual-system theories (proposed by Pinker, 1991) proposes that the lexicon and grammar are learned differently involving different cognitive abilities. The single-system theories assume that one computational system is used for all aspects of language and this system also processes all linguistic experience (e.g. non-compositional lexical items, such as 'cat', or compositional phrases, such as 'I don't know'). Unlike the dual-system theories, the single-system theories propose that lexical items and grammatical rules are learned in a similar fashion. After discussing the main features of both theories, consider the role frequency plays in these models. In single-system models frequency is considered an approximation of experience and plays a central role in developing and reinforcing linguistic units: "the more often a pattern is experienced, the easier it becomes to access and use" (Snider & Arnon, 2012:131). The mechanisms explaining the processing of frequent forms and the advantage they have differ depending on the specific single-system model but they all share the opinion that all linguistic units should be impacted by frequency effects. As for dual-system models, they view frequency effects as irrelevant (echoing generative linguistic views) because linguistic knowledge is separate from real-life probabilities

and performance that are reflected by frequency effects. This is also the reason why frequency effects are rarely discussed in these models.

Lessons on frequency in relation to the above two models can also be drawn from morphology. According to dual-system models irregular forms are stored in the mental lexicon, while regular forms are generated by the grammar. On the other hand, all forms are stored and represented by the same memory mechanisms in the single-system models. Consequently, assuming that regular forms cannot be accessed as a whole, every time an inflected form is encountered the base form (e.g. 'walk') should be activated. The frequency of the base in all of its various inflections (e.g. 'walks', 'walked', 'walking', etc.) should be reflected by the access speed. However, assuming that words are represented as a whole, the access speed is affected by the frequency of the inflected form. Frequency of inflected form has been found to be predictive of processing time, which suggests that the representation of whole-forms is possible. (Snider & Arnon, 2012)

Many studies have shown that processing is affected by two-word (i.e. bigram) frequency, as words that appear as part of more frequent bigrams are processed faster (e.g. McDonald and Shilcock, 2003) and take shorter to produce (e.g. Bell et al., 2003). However, few studies have looked at something other than bigrams, most of which have focused on highly frequent phrases and their processing, such as the study by Bell et al. (2003), which found that the ten most frequent English words are phonetically reduced when they are (more) predictable (given a previous and a following word). Snider and Arnon (2012) show that higher frequency compositional four-word phrases are processed faster: i.e. "the more often a phrase had been experienced, the faster it was processed" (Snider & Arnon, 2012:136). This shows that language users indeed possess knowledge about the frequency of phrases and words.

The question of how multiword utterances varying in frequency are represented in the mental lexicon arises. Two categories of models exist to address this question. One is the *frequency threshold approach*, which suggests independent representation of phrases with sufficient frequency in order to make their processing more efficient. However, the question remains how "sufficient frequency" is defined. The other approach proposes processing and representation are affected by every instance of usage. This approach is called the *continuous approach*. More and less frequent is thus a difference of degree. (Caldwell-Harris, Berant & Edelman, 2012)

3. The Chinese lexicon

With the rise of China's economy, not only is the country becoming an important 'global player' but the Chinese language is also becoming a valuable foreign language skill. While vocabulary learning is difficult in any foreign language, in CFL there are two additional challenges, especially for learners with an Indo-European language background: One major difficulty for learners with an Indo-European language background is that Chinese is a tone language. This means that in Chinese, as in any tone language, the pitch contour over a syllable distinguishes its meanings. The following syllables can be used for exemplification: 妈 [mā], 麻 [má], 马 [mǎ] and 骂 [mà]. These syllables are all pronounced [ma] but with different tones, which changes their meaning. The above examples' meanings are *mother*, *hemp*, *horse* and *to swear*, respectively. The different writing system is certainly another major difficulty for most learners. This study implements a written receptive vocabulary test (more on the assessment of vocabulary in chapter five), which is why the writing system will be discussed in detail later in this chapter.

After discussing the lexicon in general in the previous chapter, the focus will now turn to the Chinese lexicon, as this thesis explores vocabulary knowledge of learners of CFL. The chapter on the mental lexicon in general linguistics provided us with information on what words, multi-word units and collocations are, as well as on the organisation of the mental lexicon. When examples were used to clarify or emphasize a point, they were in English. Here features of the Chinese language relevant to this study will be discussed in a similar way to the previous chapter. What a word is in Chinese will be discussed, as well as Chinese collocations. Additionally, important factors for learning vocabulary in this language, such as the writing system, will be considered.

3.1. What is a 'word' in Chinese?

Defining what a word is in European languages is not as easy as it may seem to non-linguists but it is even more difficult to clearly define what a word is in Chinese. The reason for this being that on the one hand Chinese has far more compounds than English, for example. On the other hand it is also the case that in Chinese linguistics the term word itself is also highly controversial. For Chinese the term word cannot be defined in the same way as for flectional languages because Chinese is an isolating language with no morphological changes to the semantic units. Another difficulty lies in the fact that Chinese words also cannot be defined

orthographically-graphemically as there are no extra spaces in between the characters that belong to one word. Because spaces cannot serve for word-recognition in Chinese texts, we have to include the factor of syntactic displacement and substitutability into our definition. Based on this the following definition can be found in Guder-Manitius (2009:13): a word can be defined as “eine nicht ohne Bedeutungsveränderung trennbare syntaktisch-semantische Einheit, die mono- oder polysyllabisch (und damit auch mono- oder polygraphemisch) sein kann” (= “a syntactic-semantic unit, which is not separable without changes in meaning and can be mono- or polysyllabic (and therefore also can be mono- or polygraphemic)”). This unit is called 词 [cí] in Chinese linguistics. Interestingly, it has become a topic in Chinese linguistics only since the 20th century, while the term 字 [zì] (= ‘character’) has been prominent in Chinese linguistics for a much longer time. (San, 1998)

Although there are such difficulties in defining what a word is, there are units that can clearly be said to be words. This is the case for free morphemes, as they cannot be broken down into even smaller independent meaningful units. Bound morphemes, on the other hand, cannot be words because they cannot be used independently. Despite the fact that this seems to be pretty straightforward, it can be difficult at times to determine whether a morpheme is free or bound, which is especially true for particles. An example for an obviously bound morpheme is the plural particle 们 [men] because it can be attached to the three singular personal pronouns (我 [wǒ], 你 [nǐ] and 他 / 她 [tā]) to form the plural personal pronouns.¹ More difficult to determine is the structural particle 的 [de] because not only can it come after personal pronouns but also after practically any noun (equivalent to the English possessive ‘s’).² Certain morphemes, as the above 们 [men], can be only bound morphemes in compounds. A much larger number of morphemes can be words either on their own or in combination with another morpheme. (Guder-Manitius, 2009; Woo, 2004a)

It may be worth asking the fundamental question whether it is really necessary to identify words. In inflectional languages, it is necessary to know what a word is in order to be able to inflect correctly. In English it is necessary to know where a noun ends in order to be able to attach the

¹ Note: the plural particle 们 [men] can also be attached to certain nouns, such as 同志 [tóngzhì] (= ‘comrade’), 战友 [zhànyǒu] (= ‘comrade-in-arms’, ‘battle companion’) or 朋友 [péngyou] (= ‘friend’) to form their plurals (i.e. 同志们 [tóngzhìmen], 战友们 [zhànyǒumen] or 朋友们 [péngyǒumen]).

² However, it is important to note that the structural particle 的 [de] is not always possessive. This particle is used to express attributive structures of various kinds.

plural morpheme correctly, for example. But as the Chinese language is not an inflectional but an isolating language, it is not as imperative to search for a clear definition of the term word. In fact, by emphasizing the search for a concept of word, Chinese and its grammar will be forced into an Indo-European 'model', if you will, as has been argued by Chinese linguists (e.g. Woo, 2004a). In Chinese there is actually no difference between a 'word' and a 'phrase' when it comes to actual usage and grammatical status because there is no need to inflect them. Woo (2004a), for example, argues that concerning the Chinese language, the natural units are the syllable, the phrase-word, the clause and the sentence.

Moving on to the composition of the Chinese lexicon, it is important to highlight the fact that in China seven different dialect families exist, with the Mandarin family being the largest of them. Chen and Pasquerella (2017:31) note that about 70% of speakers belong to this dialect family. However, with great probability this number is not accurate anymore, as everyone going through the Chinese school system is forced to adopt Mandarin Chinese as an educational language. This is also true for universities that are located in autonomous areas. So these people also speak Mandarin Chinese as their educational language.

Apart from dialects or sub-standard varieties, local languages exist as well. Local languages, such as Cantonese, which in addition to its own phonetics, also has its own vocabulary and grammatical peculiarities compared to the standard language. The dialects, however, differ from Mandarin Chinese first and foremost in their pronunciation. The shared writing system leads to the dialect families having many words in common but every dialect family also has unique vocabulary and especially the basic words vary frequently. Words that are shared by the different Chinese dialects can have (slight) differences in meaning. For example, 水 [shuǐ] usually means 'water' in Mandarin, in Hakka, however, it means 'rain'. Shared vocabulary can also have interchanged meanings: the words 房子 [fángzi] and 屋子 [wūzi], for example, mean 'house' and 'room' in Mandarin Chinese, while in Cantonese 房 is 'room' and 屋 is 'house'. The differences in vocabulary can be seen from the following example taken from Woo (2004a:146) concerning the various names for the word 'maize' or 'corn' in different regional languages: 粟米 (= 'millet rice') in Cantonese, 包粟 (= 'wrapped millet') in Hakka, 游天炮 (= 'roaming sky cannon') in Fuzhounese and 玉米 [yùmǐ] (= 'jade rice') in Mandarin. However, while these terms might be more frequently used in one dialect, they can be understood or even be used in others as well. Additionally, the country's geography influenced the interaction between the dialect families. Therefore the southern dialects (such as Cantonese, Hakka or Fukienese) have more words in common with each other than with Mandarin (Chen & Pasquerella, 2017; Guder-Manitius, 2009;

Zhen, Ren & Yin, 2006) Two important features unite all the various dialects: they all share the same writing system and all of them have tones. The Chinese writing system will be discussed later in this chapter. The Chinese learned by most CFL learners also belongs to the Mandarin language family as most learners of CFL learn Standard Chinese or Pǔtōnghuà [普通话] in the People's Republic of China and Guóyǔ [国语] in Taiwan, which is based on Beijing Mandarin. For this reason the Chinese used throughout this thesis and in its empirical study is Pǔtōnghuà. (Chen & Pasquerella, 2017; Guder-Manitius, 2009; Zhen, Ren & Yin, 2006)

3.2. Word classes

Another Indo-European concept causing trouble with regard to the Chinese language is the concept of word classes. According to Woo (2004b:151) some Chinese linguists even deny the existence of this concept, while others debate on how to properly define the classes. Qi (2019) as well as Tian and Xia (2020) also agree that the classification of words causes big problems. The reason why this concept causes much controversy for Chinese linguists is quite simply that words can be used as nouns, verbs or adjectives with no change in form, thus making it difficult to distinguish words based on their formal properties (such as plural endings, for example). So, as was the case for the question “what is a word in Chinese?”, we can see here as well that it is not possible to just apply concepts that are important and necessary in inflectional languages to a non-inflectional language. In inflectional languages, different word classes call for different inflections but as in Chinese no inflection is needed, the determination of word class is rendered unnecessary. The relationship between word functions and word classes is not as regular as in other languages. While English is not a prototypical inflectional language, we can still use it as an example. In English only nouns and pronouns can function as subjects and objects in a clause. Furthermore the only word class modified by adjectives is the noun, while verbs and adjectives are qualified by adverbs. In Chinese, on the other hand, verbs, adjectives, nouns and pronouns can all function as subjects and objects. (Woo, 2004b)

Therefore, concerning the Chinese language word class should be defined in a different way by separating word class from syntactic function. A word should be examined on how it behaves in isolation not in a sentence, as is the case for Indo-European languages, for example. By implementing this method, a main classification will be found for most words. The basic division of words is between *content* (lexical) words and *function* (grammatical) words, which were already discussed in general in the previous chapter (2. *The lexicon in general linguistics*). Chinese linguists (e.g. Qi, 2019) call into question whether function words and content words

can and should be separated in Chinese. This is due to the fact that whether a “word” is actually an individual word or only part of a word and to which word class it belongs to can only be clearly distinguished from its context and not in isolation. If we take the word 生产 [shēngchǎn] whether it means ‘production’ or ‘to produce’ only becomes apparent from its surroundings and its function in the sentence. However, I will argue that the distinction is at least a helpful construct in creating useful subdivisions in the Chinese lexicon in order to better describe the test conducted for this study in chapter six on research methodology.

3.2.1. Content words

The words of all languages can be grouped into classes with different semantic and syntactic functions. On the syntactic level, there is no universal definition of word classes because morpho-syntactic structure is language-specific (this includes the division into word classes). On the semantic level, on the other hand, Gärdenfors (2014:181) points out that certain universal and prototypical structures among words can be identified. These prototypical structures include nouns, verbs and adjectives, which are typological prototypes because they are language-universal, meaning they occur in almost all languages. These three categories denote different prototypical semantics: nouns denote entities, while verbs denote processes and adjectives denote qualities. Each language, however, has its own criteria for identifying these and other categories depending on its syntactic structure. The criterion for identifying Chinese content words, according to Woo (2004b:161) is rather simple. Nouns, pronouns, verbs and adjectives can all be used to modify nouns. Therefore, the possessive particle 的 [de] can follow them immediately.

We will now take a closer look at each of these categories of content words, starting with *nouns*. On a formal syntactic level, while English or German nouns can be identified by checking whether an article can precede a specific word, this is not possible for Chinese nouns, as articles do not exist. But nouns can be identified in another way. As mentioned above, adjectives are linked to nouns by the structural particle 的 [de]. Consequently this structure can be used to identify nouns, as nouns can be defined as units that can immediately follow this particle. This is true for all common nouns; whether they are animate, inanimate, abstract or non-existent does not play a role in this case. Semantically speaking English and Chinese nouns are similar in that they denote entities, but English can form abstract nouns more easily by suffixing. In English, abstract entities are typically denoted by nouns created by suffixing, while in Chinese abstract nouns were usually formed by coordinate compounds, such as 大小

[dàxiǎo] (=‘size’, literally meaning ‘big-small’), at least until the 20th century. In the early 20th century a large number of nouns were also formed by adding suffixes, such as 性 [xìng] (= ‘property’) or 者 [zhě] (=‘person’). Categorical nouns, for example are compounds with one morpheme denoting its category. This commonly applies to fish, flowers and trees, for example. The following compounds denoting different types of fish all have 鱼 [yú] (= ‘fish’) as their last component: 鲨鱼 [shāyú] (= ‘shark’), 鳗鱼 [mányú] (= ‘eel’), 鲤鱼 [lǐyú] (= ‘carp’) and 墨鱼 [mòyú] (= ‘cuttlefish’). While ‘cuttlefish’ is an example for a categorical noun, this type of noun is not that common in English. (Woo, 2004b)

Next we will examine Chinese *pronouns*. As in English, they can be divided into three types: *personal*, *demonstrative* and *interrogative*. The personal pronouns are similar to the English ones in that there are pronouns for first, second and third persons singular and plural but there are certain special features, which should be shortly mentioned here. The plural personal pronouns, as we mentioned before, are formed by adding the plural particle 们 [men] to the singular forms, i.e. by adding the particle to the first person singular 我 [wǒ] (= ‘I’) one gets the first person plural 我们 [wǒmen] (= ‘we’). In Chinese one can also find a polite second person singular form 您 [nín], which is similar in function to the German ‘Sie’. Secondly, as for demonstrative pronouns, in Chinese different words exist that express proximity or distance. These are 这 [zhè] (= ‘this’) and 那 [nà] (= ‘that’). Thirdly, Mandarin Chinese has quite a number of interrogative pronouns, more than other dialects of Chinese, these include 谁 [shéi] (= ‘who’), 什么 [shénme] (= ‘what’) and 哪 [nǎ] (= ‘which, what’), for example.³

Another difference between Chinese and English or German concerns the characterisation of *verbs* because Chinese verbs do not have tense. One of the main characteristics of Chinese verbs is that they can be identified by their position in a sentence, as they mainly function as predicates. So how can we identify Chinese verbs? One way is by using adverbs of manner because in Chinese they do not qualify adjectives. Adverbs of manner, similar to possessive relations, are also indicated by a particle, which in this case is 地 [de]. Therefore, verbs can be described as words that are preceded by an adverbial particle, which is actually parallel to the English adverbial suffix *-ly*. But in some cases identifying verbs this way can be difficult, especially for verbs such as 有 [yǒu] (= ‘have’) and 是 [shì] (= ‘be’) because it is hard to think of

³ In Written Chinese additional pronouns are in use, which are derived from the language used in pre-modern written language, such as 此 [cǐ] (= ‘this’) and 其 [qí] (= ‘his, her, its, their, that’).

an adverbial of manner to qualify verbs not involving actions. This kind of verbs can be identified by adverbs of time instead of adverbs of manner, such as 快 [kuài] (= 'about to'). Furthermore, verbs can be divided into verbs with and verbs without objects. However, verbs are the most complex of the word classes. Because besides 'simple' verbs that resemble English verbs, compound verbs can be found as well. According to Woo (2004b:174) these can be divided into three different types: verb-object, verb-complement and subject-predicate compounds. *Verb-object compounds* (离合词 [líhécí]) consist, as the term suggests, of a verb and an object, seemingly acting as a single verb. 见面 [jiànmiàn] (= 'to meet') and 睡觉 [shuìjiào] (= 'to sleep') are examples of such verb-object compounds. *Subject-predicate compounds* can be also used as predicates. This type of compound can be best exemplified with a sentence (Note: the subject-predicate is underlined in the examples and their translations): 我头疼 [wǒ tóuténg] (= 'I (have a) headache') or 他嘴笨 [tā zuǐ bèn] (= 'He (is) inarticulate'). The *verb-complement compound verbs* can be divided into different types, two of which are particularly important. First, there are compound verbs where the complement denotes directions (趋向补语 [qūxiàng bǔyǔ]), such as *up, across, over* and the like. The following two verbs are examples of this type: 进来 [jìnlái] (= 'come in') and 上去 [shàngqù] (= 'go up').⁴ The second type that should be mentioned here, are verbs with complements of results (结果补语 [jiéguǒbǔyǔ]), where the complement describes the result of the action. Examples of this type of complement are 听到 [tīngdào] (= 'to hear', 听 [tīng] on its own has the meaning 'to listen') and 写对 [xiěduì] (= 'to write something correctly'). However, it should be noted that verbs are not the only word class that can be combined with complements, as adjectives can behave in the same manner. (Woo, 2004b; Zhu, 2007)

The third main word class are *adjectives*, which according to Woo (2004b:175) some Chinese linguists consider to be verbs. This is because Chinese adjectives differ from English or German adjectives in one fundamental way: in combinations with the adverb 很 [hěn] they can function as the predicate of a sentence on their own, without a verb, where German and English would

⁴ It should be noted that two types of complements of direction exist: the *simple complement of direction* and the *compound complement of direction*. The former denotes the direction of an act in reference to the speaker; the examples above are then such simple complements of direction. The compound complements of direction can be used in a similar manner, but they give more detail about the direction of an action. Examples of the compound complement of direction are 站起来 [zhànqǐlái] (= 'stand up') or 跑出去 [pàochūqù] (= 'run out').

use the verb *sein* / *to be* combined with a relational adjective. The following sentences contain an adjective as their predicate (Note: the predicate is underlined in the examples and their translations): 德语很难 [Déyǔ hěn nán] (= ‘German [is] difficult’) or 天空很蓝 [tiānkōng hěn lán] (= ‘the sky [is] blue’). With adjectives we can distinguish two different types. On the one hand there are adjectives that can be qualified by adverbs of degree such as 最 [zuì] (= ‘most’). When using this method of identifying adjectives, it is important to keep in mind to apply the adverb of degree only to the adjective on its own, not to combinations containing that word. On the other hand some adjectives cannot be identified by applying adverbs of degree. These are non-comparable adjectives, such as 共同 [gòngtóng] (= ‘common’), or adjectives that are part of a set, such as 一年生 [yīniánshēng] (= ‘annual’). These can be identified as adjectives only with great effort. A short definition of Chinese adjectives is that an adjective describes a noun in (more) detail without having to form a phrase to do so. (Woo, 2004b)

It has been shown above that the different classes of content words, such as nouns, verbs and adjectives, can be established in Chinese as well. However, Chinese linguists (e.g. Qi, 2019) argue over some of these distinctions. One such example would be whether adjectives should be distinguished from verbs or whether they actually are verbs. The reason for this is that adjectives can function as predicates on their own without a verb, as was shown above.

Two additional important issues in distinguishing word classes in Chinese should also be discussed here. One main issue is that words can have the grammatical characteristics of different word classes depending on their surroundings or context of use and thus can belong to different word classes in different contexts without changing their form. Another important issue is that of nominalisation. The traditional view on nominalisation is that when verbs or adjectives function as subjects or objects of a sentence, they are nominalised. 游泳 [yóuyóǒng] (= ‘swim’) in 我喜欢游泳 [wǒ xǐhuan yóuyóǒng] (= ‘I like swimming’) would be an example of this view on nominalisation. Zhu et al. (1961) and other Chinese linguists oppose this view because if all verbs and adjectives functioning as subjects or objects of a sentence are regarded as nouns, then words could not be classified at all. Later on, Hu and Fan (1994) distinguished two different types of nominalisation: ‘nounisation’ (名词化 [míngcíhuà]) at the syntactic level and nominalisation (名物化 [míngwùhuà]) at the semantic level. Guo (2002) also suggested that all words should be divided into word classes based on a lexical and syntactic level. If we consider the above example 游泳 [yóuyóǒng] (= ‘swim’) again, it can be said that at the semantic level it is a verb and that it can change word class at the syntactic level based on its surroundings and

role in the clause. Furthermore, a change in word class can also mean a change in meaning: the word 派 [pài], for example, as a noun means ‘school [of thought, of art]’ or ‘faction’, while as a verb it takes on the meaning ‘[to] send’ or ‘[to] dispatch’. All these issues cause controversy and continuing debate around the distinction of word classes in Chinese. (Liao, 2019; Qi, 2019; Tian & Xia, 2020)

3.2.2. Function words

The second major word-category, function words, will be discussed more briefly, as they do not play a role in the empirical study conducted for this thesis (the word classes important for the empirical study are part of the content words category, i.e. adjectives and nouns). First, the categories of Chinese function words, which are similar to English or German, will be discussed. One of these is *numbers*. Chinese uses a decimal numeration system, which means that ‘twelve’ is ‘ten two’ (i.e. 十二 [shíèr]) in Chinese. The second category is *adverbs* and *conjunctions*, which are similar to their English or German counterparts. One important difference is, however, that Chinese conjunctions are often used in pairs, such as 因为...所以... [yīnwèi...suǒyǐ...] (= because...therefore...) or 虽然...但是... [suīrán...dànshì...] (= although...but...). (Woo, 2004b)

Although there are some function words in English, German and Chinese that are similar to each other, it should be noted that certain function words have no counterparts, for example English and German do not have particles, while Chinese does not have articles.

One prominent category of function words in which English or German and Chinese differ is that of *measure words*. This type of word is an obligatory element positioned between a number and a noun. Although measure words usually refer to shape or composition of the quantified objects, there are also abstract measure words to be found. The most general measure word is 个 [gè], which can be used with persons, but also with objects without a distinguishing form. This is a major difference between English or German and Chinese because although the first two languages do have some measure words: *cup of coffee*, *piece of cake* and *bottle of perfume* are examples of English and *ein Laib (Brot)* (= ‘a loaf [of] bread’), *ein Glas (Wasser)* (= ‘a glass [of] beer’) and *eine Prise (Salz)* (= ‘a dash [of] salt’) are examples of German measure words. However, the number of English and German measure words is relatively small compared to Chinese. The reason for this is that in English and German nouns can be quantified with a

number without having to add a measure word in most cases, while in Chinese it is obligatory for a measure word to precede a noun that is quantified with a number. (Woo, 2004b)

Another category that differs quite radically from English or German is the one of *particles*. These denote certain grammatical functions. Particles can be divided into three types. First, particles used for grammatical structure and for indicating meanings that in inflectional or agglutinating languages are expressed by bound morphemes (e.g. case endings, tenses, aspectual endings), such as the possessive particle 的 [de] and the particle 地 [de] indicating adverbials, which were mentioned before. Second, particles can also be used for marking aspects, such as progressive or perfective aspects. In Mandarin Chinese these two aspects are expressed by the particles 着 [zhé] and 了 [le] respectively. Third, Chinese also uses intonation particles. Intonation particles are placed at the end of a sentence in order to establish the tone and the type of the sentence. The following are examples of such intonation particles: 啊 [a] to express enthusiasm or impatience, 呢 [ne] for rhetorical questions or 吧 [ba] to indicate suggestions, requests (or a mild command), as well as to indicate consent or approval. (Woo, 2004b; Woo, 2004c)

As a concluding remark, the possibility of content words changing into function words over time should be mentioned. This process of grammaticalisation was undergone by most Chinese prepositions and conjunctions. They started out as verbs, which gradually underwent semantic bleaching, and finally turned into function words. Measure words on the other hand often started out as nouns. (Woo, 2004b)

3.3. Collocations in the Chinese lexicon

Which words can form collocations with each other always depends on the language and culturally established thought patterns among its speakers. While a general definition of collocations was given in the previous chapter, the focus will now turn to collocations in the Chinese language. The first thing to note is that research on Chinese collocations seems to be a rather neglected area, which was also noted by Liu (2001). Nevertheless, this section will attempt to summarize the research that exists on finding a definition and types of Chinese collocations.

Hu and Xiao (2019) report that, similar to collocations in general (see chapter 2), Chinese collocations were defined by Xu et al. (2009) as recurring combinations of two or more content words with grammatical and semantic connections between them. This definition considers

frequency of co-occurrence, while also stressing syntactic and semantic relationships between the words. A similar definition can be found in Zhen, Ren and Yin (2006) and Li (2008) who say that: collocations are ‘habitual’ co-occurrences of words.

However, such a definition also has two main deficiencies according to Hu and Xiao (2019): on the one hand it limits collocations to content words because it does not take in to account the importance of function words, such as prepositions, auxiliary words or conjunctions. On the other hand it considers only cross-language word classes (e.g. predicate and object or modifier and noun) but ignores other syntactic and semantic relationships important in Chinese, such as preposition and locator or measure word and noun. Thus, Hu and Xiao (2019:137) suggest a slightly altered definition of Chinese collocations: a collocation is formed by two or more words that occur together frequently, these words can be adjacent and non-adjacent, content and function words and must have a syntactic or semantic relationship.

After defining Chinese collocations, which is similar to the general definition of collocations but has been extended to cover also function words by Hu and Xiao (2019), we need to consider whether different types of collocations exist. According to Zhen, Ren and Yin (2006), in Chinese academic circles this question of how to categorise word collocations is one of the main research-foci in this area. In a first step Chinese collocations can be divided into two categories: *fixed collocations* (固定搭配 [gùdìng dāpèi]) and *free collocations* (自由搭配 [zìyóu dāpèi]). The former are inseparable word combinations, as the name ‘fixed collocations’ suggests. This means that it is not possible to divide the words or add something in between them. Five types of expressions belong to this category, all of which will be briefly discussed here. First, there are the Chinese *idioms* (成语 [chéngyǔ]), which are fixed combinations making use of written language (书面语 [shūmiànyǔ]) whose structural components cannot be changed or reversed at will. Structurally most Chinese idioms are composed of four characters with attention being paid to harmony of the syllables and tones. Their origins lie in myths, legends and allegorical stories. Examples for Chinese idioms are 守株待兔 [shǒuzhū dàitù] (which literally means to guard a tree stump waiting for rabbits with the metaphorical meaning ‘to wait idly for opportunities’ or ‘to trust chance rather than to show initiative’) or 空中楼阁 [kōngzhōng lóugé] (which literally means ‘pavilion in the air’ and has the metaphorical meaning ‘imaginary future plans’; this Chinese idiom can also be compared to the German ‘Luftschlösser’).

Second, Chinese *proverbs* (谚语 [Yànyǔ]) also belong to this category, as they are also fixed combinations but unlike the idioms they make use of spoken and more simple language.

Proverbs reflect certain (social) principals and laws, such as 一寸光阴一寸金 [yī cùn guāngyīn yī cùn jīn], which is similar to the phrase ‘time is money’ in English, as it literally means an interval of time is worth an ounce of gold. The following is another example for a Chinese proverb: 上梁不正下梁歪 [shàngliáng bù zhèng xiàliáng wāi]. This phrase literally means if the upper beam is not straight, the lower beam will be crooked. The figurative meaning can be taken to be ‘subordinates imitate their superior’s vices’. A similar saying can be found in German: ‘der Apfel fällt nicht weit vom Stamm’ (= ‘the apple does not fall far from the tree’), which refers to family members, especially parents and their children, being similar to one another.

Allegorical sayings (歇后语 [xiēhòuyǔ]) are the third type of fixed collocation. This type of saying has its own special structure: the first part is descriptive but is similar to ‘a riddle in the riddle’, while the second part carries the message of the saying and is so to speak ‘the answer to the riddle’. 泥菩萨过河，自身难保 [nípúsà guò hé, zìshēn nán bǎo] is one example of such a saying, meaning ‘like a clay idol fording a river – unable to save oneself (let alone others)’. The following saying literally meaning ‘the weasel goes to pay his respects to the hen – without good intentions’ is another example: 黄鼠狼给鸡拜年，没安好心 [huángshǔláng gěi jī bàinián, méi ān hǎoxīn].

The fourth type of fixed collocations are *aphorisms* (格言 [géyán]). These are sayings with incisive content containing admonishments. 千里之堤，溃于蚁穴 [qiānlǐ zhī dī, kuì yú yǐxué], which literally means ‘one ant hole may cause the collapse of a great dike’ and figuratively means that huge damage may result from a moment’s negligence. A similar saying in English would be ‘the straw that broke the camel’s neck’ and in German ‘der Tropfen, der das Fass zum Überlaufen bringt’ (= ‘the drop that causes the barrel to overflow’). 冰冻三尺，非一日之寒 [bīng dòng sān chǐ, fēi yī rì zhī hán] is another example for a Chinese aphorism, which literally means ‘it takes more than one cold day for the river to freeze three chi deep’ and figuratively has the meaning ‘trouble has been brewing for some time’ or ‘problems have piled up’, are examples for Chinese aphorisms.

The fifth and last type of fixed collocations can be called *colloquial expressions* (惯用语 [guànyòngyǔ]). This type of expression does not have a fixed number of words, usually has a metaphorical meaning and uses spoken language. The following expressions or phrases can exemplify this: 炒鱿鱼 [chǎo yóuyú] (= ‘to fire somebody’, the literal meaning is ‘fried squid’), 穿小鞋 [chuān xiǎoxié] (= ‘to make life difficult for somebody’, the literal meaning is ‘to make

somebody wear tight/small shoes’) and 车到山前必有路 [chē dào shān qián bì yǒu lù] (= ‘everything will turn out for the best’, literally meaning ‘when the car will get to the mountain, there will be a way around [it]’). (Liu, 2001)

After focussing on the fixed collocations, our attention will now turn to *free collocations* in Chinese. Unlike fixed collocations, this type of collocation is made up of flexible word combinations but still adhering to grammatical structure and semantic connection of the words. These words often appear in sentences together, although they are not fixed in form. Content words, such as nouns, verbs and adjectives, can be used to form such collocations. We will focus on three types of free collocations: noun-based, verb-based and adjective-based collocations. The first type of free collocation that is going to be discussed is *noun-based collocations* (以名词为基词的搭配 [yǐ míngcí wèi jīcí de dāpèi]), which have a noun as their base word. Because nouns denote entities, their qualities are often described with adjectives. Thus, noun-based collocations often consist of an adjective and a noun, such as 高山 [gāo shān] (= ‘high mountain[s]’), 旧社会 [jiù shèhuì] (= ‘old society’) and 国际贸易 [guójí màoùyì] (= ‘international trade’). Liu (2001:57) also notes that certain Chinese words have strong collocation ‘ability’ or potential and can form many collocations. One such word is 局 [jú] (= ‘bureau, office’), which can be combined with other related terms to form the name of institutions, such as 政治局 [zhèngzhìjú] (= ‘politburo’) or 外事局 [wàishìjú] (= ‘foreign affairs bureau’). However, 局 [jú] can also form phrases consisting of a modifier and the word it modifies (偏正词组 [piānzhèng cízǔ]), such as 败局 [bàijú] (= ‘lost game’, ‘lost battle’) or 饭局 [fànjú] (= ‘dinner party’, ‘feast’). (Liu, 2001)

Verb-based collocations (以动词为基词的搭配 [yǐ dòngcí wèi jīcí de dāpèi]) are the second type of free collocations. Because the main syntactic function of a verb is to act as a predicate, this type of collocation usually consists of either a verb and an adverb (as an adverbial) or a verb and a noun (as an object), such as 反复琢磨 [fǎnfù zhuómó] (= ‘reflecting repeatedly’) or 实现现代化 [shíxiàn xiàndàihuà] (= ‘realising modernisation’). Some verbs also have a strong collocation potential and can be combined with many nouns or adjectives to form different collocations. One verb that is commonly used to form collocations is 打 [dǎ], which can form a large number of collocations, such as 打电话 [dǎ diànhuà] (= ‘to make a phone call’), 打工 [dǎgōng] (= ‘to work a temporary job’), 打牌 [dǎpái] (= ‘to play cards’) or 打官司 [dǎ guānsi] (= ‘to go to court’, ‘to engage in a lawsuit’). (Liu, 2001)

The third and last type of free collocation to be discussed here is the *adjective-based collocation* (以形容词为基词的搭配 [yǐ xíngróngcí wèi jīcí de dāpèi]). Because adjectives describe qualities, most of them can be modified with adverbs of degree, such as ‘very’ or ‘most’, to form adjectival phrases. Additionally, adjectives can also be paired with other adjectives or reduplicated to express a more vivid description of something, as in the following two examples: 白花花 [báihuāhuā] (= ‘shining white’) and 金灿灿 [jīncàncàn] (= ‘golden’, ‘glittering’). Some adjective collocations can appear similar to the four-character idioms, discussed above. This is the case when two disyllabic adjectives are combined. However, they do not have a fixed form, also if they are paired together often. 简单易行 [jiǎndān yìxíng] (= ‘simple [and] easy’), 热情周到 [rèqíng zhōudào] (= ‘warm-hearted [and] considerate’) and 清脆悦耳 [qīngcuì yuè’ěr] (= ‘melodious [and] pleasant to the ear’) are examples of such four character adjective collocations. (Liu, 2001) In addition to these types of free collocations, numerals and conjunctions and other word classes can also form collocations, which will be discussed below in connection with the categorisation of collocations by Hu and Xiao (2019), which they call “Chinese-specific collocations”.

Hu and Xiao (2019:137-140) go about categorising Chinese collocations in a different way, in consonance with their definition described earlier in this section. They do not distinguish between fixed and free collocations, but between (*cross-language*) *universal collocations* and *Chinese-specific collocations*. The former category consists of four different types and the latter of five different types of collocations. Cross-language universal collocations (跨语言通用搭配 [kuà yǔyán tōngyòng dāpèi]) are types of collocations that exist in most, if not all, languages. This includes verb and object, subject and predicate, attribute and a word and adverb and a word. As for the Chinese-specific collocations (汉语特有搭配 [Hànyǔ tèyǒu dāpèi]) the five different types are: measure word and noun, “frame structure preposition collocation”, preposition and verb, collocations with complements and collocations with conjunctions. We will look at each of these in more detail: first consider *measure words and nouns*. Measure words, as discussed in the section on word classes, are obligatory words between numerals and demonstratives and nouns. According to Hu and Xiao (2019) the co-occurrence of measure words and nouns can be seen as a collocation, as a large number of nouns and verbs require specific measure words. Additionally, often a latent semantic or cognitive connection between the nominal measure words and the modified noun can be found. Consider the measure word 条 [tiào] for example: generally, it is used to define long, narrow or thin objects, such as rivers (河 [hé]), roads (路 [lù]) or fish (鱼 [yú]).

Second, the “*frame structure preposition collocations*” (框式介词搭配 [kuāngshì jiècí dāpèi]) are another type of Chinese-specific collocation. Chinese contains a high number of prepositional frame structures, such as 在…上 [zài...shàng] (= ‘on’) or 自…以来 [zì...yǐlái] (= ‘since’). Hu and Xiao (2019) define these structures as combinations of prepositions and postpositions. Said postpositions complement the preposition semantically and grammatically. Generally speaking, these frame structure prepositions represent time, place, manner and purpose and serve as components to modify predicate or noun phrases.

Third, collocations of prepositions and verbs exist as well, which are different from the frame structure prepositions discussed before, as they do not require postpositions to form prepositional phrases. This type of prepositions mainly represents the agent or the recipient of an action, such as the preposition 把 [bǎ], for example. This preposition is used to shift the object ahead of the predicate, which has to be reduplicated or accompanied by complements: e.g. 把衣服洗得干干净净 [bǎ yīfu xǐ dé gāngān jìngjìng] (= ‘wash the clothes clean’).

The fourth type of Chinese-specific collocations is collocations with complements. Complements always follow the predicate and supplement the description of said predicate (i.e. direction, degree, result, etc.). Because the predicate restricts the components that can act as a complement, Hu and Xiao (2019) define the combination of predicate and complement as a unique type of Chinese collocations.

The fifth and last type of collocation specific to Chinese are collocations with conjunctions. Both adverbs and conjunctions can connect words, phrases or clauses and as mentioned before these are often used in pairs, as mentioned earlier in the section on word classes. Hu and Xiao also define these fixed pairs of conjunctions as a type of Chinese collocations.

However, the author of this study considers these types of “Chinese-specific collocations” distinguished by Hu and Xiao (2019) to represent the category of *grammatical collocations*, which can be found in Benson, Benson and Ilson (2009:XIII-XIV), who distinguished between grammatical and lexical collocations. According to Benson, Benson and Ilson (2009) a grammatical collocation consists of a dominant word (i.e. usually noun, verb or adjective) and a preposition or a grammatical structure. An example provided by Benson, Benson and Ilson (2009:XIII) in both English and German of a collocation of a noun and a preposition is ‘hostility towards’ and ‘Feindseligkeit gegenüber’. Therefore the second and third type of “Chinese-specific collocations” described above definitely fall into this category of grammatical

collocations. Additionally, it can also be argued that the first, fourth and fifth type described above (measure word and noun/verb, collocations with complements and collocations with conjunctions) also fall into this category of grammatical collocations, as these can be categorised as “grammatical structure”.

Thus, it can be seen that the approaches to categorise collocations suggested by Liu (2001) and Hu and Xiao (2019) are quite different. This thesis will make use of Liu’s categorisation because the distinction between fixed and free collocations is more applicable to the present study.

3.4. The Chinese writing system

While writing can be seen as a medium to store linguistic utterances or texts which overcomes space and time, it also serves as a way of communication and so plays an important role in learning any foreign language. For European language learners of CFL the Chinese writing system is a big challenge, as the difference between the English or German and the Chinese writing systems seems to be quite enormous. The Chinese writing system does not represent language on a phonemic but on a syllabic level, which is different from the English or German writing system, where an alphabetical script represents a language (more or less) phoneme by phoneme. The Chinese characters rather correspond to spoken syllables and remain graphically unchanged, as the syntactic function of a word has no influence whatsoever on a character’s inner structure and written realisation. The grammatical function of a word is therefore solely evident through word order, context or other markers. This writing system has two major advantages compared to other systems. One advantage is that through this way of representing syllables it is possible to generate new polysyllabic words made up of already known ones, which makes it easier to learn certain new words than it is to learn new words in alphabetic languages. However, this facilitation is only true if the learner has already acquired at least 500 Chinese characters. Prior to this, learning new vocabulary is more difficult than when learning another alphabetic language, as not only the meaning of a word has to be learned but also its pronunciation, which – in contrast to English or German – is not contained in the orthography. The second advantage is that this writing system secures understanding across dialects in China, as it is mostly independent of articulation. Compared to that, alphabetic writing systems adapt to the regional variations of articulation at least to a certain extent. (Chen & Pasquarella, 2017; Guder-Manitius, 2009)

In order to better understand the challenges learners of CFL encounter when learning vocabulary that will be discussed in the next chapter on vocabulary learning, the structure and composition of Chinese characters will be discussed here shortly. The building blocks of the Chinese characters are strokes, which do not, however, represent phonological or meaning units. Therefore the correspondence between speech and the characters is relatively weak. How then did the Chinese characters form? Generally six different principles of character formation exist: the first principle being that characters emerged from the illustration of their reference objects, wherefore this type of characters is called pictogram (象形 [xiàngxíng]). But this does not mean that we can still see the meaning in the character's graphic form today. Examples for pictographic characters are 山 [shān] (= 'mountain'), 日 [rì] (= 'sun') or 雨 [yǔ] (= 'rain'). Secondly, there are characters, which represent relational concepts. These are called indicators (指事 [zhǐshì]) and mostly are a pictogram with one indicative stroke to label a certain aspect of said pictogram. The characters 三 [sān] (= 'three'), 上 [shàng] (= 'on, above') or 本 [běn] (= 'root') are examples of this principle of formation.

Next we have meaning compound characters (会意 [huìyì]). Characters formed by this principle are composed of two or more other characters and their meaning can be gathered from the semantics of the individual components – at least to a certain extent. The following characters are examples for this principle: 明 [míng] (= 'bright'; formed by combining 日 [rì] 'sun' and 月 [yuè] 'moon'), 信 [xìn] (= 'believe'; formed by combining 人 [rén] 'person' and 言 [yán] 'words') or 林 [lín] (= 'forest'; formed by two 木 [mù] 'tree').

Another principle of character formation is the semantic-phonetic compound character (形声 [xíngshēng]), which as the name suggests is made up of a component with reference to the character's meaning and a component with a phonetic reference to its pronunciation. For example, the character 同 [tóng] is a component of the following four characters each sounding similar to it: 铜 [tóng], 酮 [tóng], 筒 [tǒng] and 洞 [dòng]. So 同 [tóng] is the phonetic component in these four characters. Unfortunately, due to changes over many centuries in a great number of cases the pronunciation of characters of this type is no longer identical with the pronunciation of its phonetic component. As for the semantic component of characters we can look at the following five characters as examples: 腿 [tuǐ], 脚 [jiǎo], 肘 [zhǒu], 脑 [nǎo] and 肚 [dù]. They all share a common semantic component 月 [ròu] (= "flesh")⁵. But this component gives us only a

⁵ 月 [ròu] is the radical form of the character 肉 [ròu], which means 'flesh' or 'meat'.

hint about the actual meaning of these characters, as these characters mean “leg”, “foot”, “elbow”, “brain” and “stomach”, respectively. So the radical 月 [ròu] tells us the general semantic area of the characters used in this example, which in this case is parts of the body. What can be said in general is that the semantic components give us only loose meaning boundaries, which in itself can already be helpful when trying to decipher the meaning of an unknown character.

The fifth principle is the principle of phonetic borrowing of characters with the same pronunciation (假借 [jiǎjiě]). This principle does not generate new characters but gives a character a new meaning. This principle was especially applied to words without tangible content, which could not be graphically represented in an easy way. The character 其 [qí] (= ‘his, her, its, their’) is a good example for this principle, which originally had the meaning of ‘shovel’ or ‘dustpan’. The character 要 [yào] (= ‘want’) is another good example, which used to mean ‘waist, small of the back’.⁶

The last principle of character formation is also the most controversial. This is the principle of mutually defining characters (转注 [zhuǎnzhù]). According to Guder-Manitius (2009:32) a few different definitions can be found but what they all have in common is that the form, pronunciation or meaning are not of importance but that it defines itself via comparing two characters. One example are the characters 考 [kǎo] and 老 [lǎo], which share similarities in both their graphic representation and their pronunciation. Other characters, which can also be attributed to this category are 后 [hòu] and 司[sī], as they practically are mirror-images of each other. (Chen & Pasquarella, 2017)

As can be seen above, the pronunciation of characters is not apparent from their form. This is the reason why various systems of transcription have been established. The most common one today, especially for learners of CFL, is Hànyǔ Pīnyīn [汉语拼音] because it shows the pronunciation of Pǔtōnghuà. Pīnyīn uses Latin letters or combinations of letters (*zh-*, *ch-*, *sh-*, *_ng*) and their articulation is clear in most of the cases because Pīnyīn uses a so-called flat orthography. Of course there are exceptions to this rule (the graphemes *a*, *e*, *u* and *i*) because of different realisation in different phonemic environments. But as with any attempt of alphabetisation of the Chinese language and although it has a fixed way of writing, Pīnyīn also

⁶ Today the character for ‘shovel’ or ‘dustpan’ is 箕 [jī], written with the semantic component ‘bamboo’ and the character for ‘waist, small of the back’ is 腰 [yāo], written with the semantic component ‘flesh’.

has the problem of defining word boundaries, since word boundaries are not always clear due to the fact that there are no spaces between characters in written text (i.e. also no spaces between characters belonging to different words). (Guder-Manitius, 2009)

A very popular question is how many characters the Chinese writing system contains. It is not easy to find an answer to this question. One reason for this is that the beginnings of the writing system and the differentiation between graphic variants and variants with different meanings are not clear. Character lexicons compiled at different times contain different numbers of characters. The earliest known character lexicon (121 AD) contains 9.353 characters, while the biggest lexicon of the imperial era, the *Kāngxī Zìdiǎn* [康熙字典] (1716), contains about 47.000 characters, for example. The major reference works of our time reach numbers of about 50.000 characters as well. But these huge compilations also contain “dead characters”, which are no longer in use in today’s Chinese language. Another point of much discussion is the number of characters needed to read Chinese texts fluently. The number of characters that are being used today is often declared to be ten to twenty thousand but in reality this is much too high. According to literature (e.g. Guder-Manitius, 2009:45) children in the People’s Republic of China learn 2.500 to 2.800 characters in six years of school. In analogy, it makes sense to also restrict the number of characters aimed for in teaching CFL. But of course not only the number of characters is important; their frequency of occurrence plays a major role as well. More on frequency of occurrence will be discussed in chapter four on the learning of vocabulary. (Guder-Manitius, 2009)

4. Learning vocabulary

On an intuitive level, the term *vocabulary learning* is pretty clear to most people who have learned a foreign language: it means learning the many words you did not know before. One important notion in relation to word knowledge is the notion of *learning burden*. This notion describes the amount of effort it takes to learn a word. Each aspect of word knowledge can contribute to the word’s learning burden. But the more familiar learners are with the patterns and knowledge represented by a word, the lighter its learning burden. This knowledge can come from either the learner’s L1, other languages learned or prior knowledge of the language in question. Another factor to lighten the learning burden is the relationship of the languages: the closer the languages are related, the lighter the learning burden. But if first language and target language are very distant, how can the teacher help to lighten the learning burden? This

can be done by highlighting systematic patterns and analogies within the L2 or foreign language and by also drawing attention to connections to the L1. (Nation, 2013) Now consider the learning burden of CFL. Two factors that impact the learning burden of Chinese have already been mentioned at the beginning of the previous chapter: the Chinese writing system and the tones. Additionally, the learning burden of CFL is impacted by the fact that English and German belong to a totally different language family than Chinese (Indo-European vs. Sino-Tibetan language families). Thus, the likelihood for European learners of Chinese to be able to draw on L1 knowledge or previously learned languages that can function as support for CFL vocabulary learning is very low indeed.

Although, the term vocabulary is known intuitively by most people, when dealing with the topic of vocabulary from a scientific point of view, we need to look at different aspects that play a role in learning vocabulary, such as different dimensions of vocabulary knowledge or what knowing a word actually means, which will be the focus of this chapter.

4.1. What does it mean to ‘know a word’?

To answer the question posed in the section heading, it is important to remember that words are not isolated language units but are part of many interrelated systems and levels, which is why there are many different degrees of knowing a word. Different aspects or degrees of knowing a word exist, with an important distinction being the one of *receptive and productive word knowledge*. Receptive word knowledge describes the idea of receiving and understanding language input produced by others, which can be divided into spoken and written input modes. Productive word knowledge on the other hand describes the speaker’s active use of words in written or spoken production. Different aspects of cognition are included in receptive and productive language use. The former includes perceiving a word form and then retrieving its meaning (while listening or reading), while the latter includes the wish to express a certain meaning (through retrieving and producing the appropriate spoken or written forms). Another distinction between productive and receptive vocabulary is their way of activation. Productive vocabulary is activated by other words and therefore includes incoming and outgoing links between words. Receptive vocabulary items on the other hand are only activated by external stimuli (i.e. hearing or seeing their forms) (Meara, 1990). Some researchers (e.g. Nation, 2013) share the opinion that receptive vocabulary actually includes productive vocabulary, as well as three other kinds of vocabulary, which are partly known words, low-frequency words not readily available for use and words which are avoided in productive use. To a certain degree these three categories overlap. A problem with this position is that it is strongly based on the aspect of

use and not the aspect of knowledge (Nation, 2013:48). Now concerning the scope of this receptive and productive distinction, it is true that it applies to a variety of knowledge and use of language. What many researchers agree on is that receptive word use/knowledge is less difficult than productive use. There are generally four explanations for this phenomenon, which do not compete with but complement each other (Nation, 2013:51-52). First, “productive learning is more difficult because it requires an extra learning of new spoken or written output patterns” (Nation, 2013:51). This is especially true for languages that use different writing systems from the learner’s L1 (e.g. a learner of CFL with English or German as his/her L1) and (some) different sounds/sound combinations. The reason for this is that the knowledge of word form needs to be more precise for productive use, while learners only need to know a few distinctive features of the word form for receptive use. A second factor involved is the fact that receptive vocabulary use gets more practice than its productive counterpart, at least in an average foreign language-learning situation. Thirdly, the access to words also plays an important role. At the early stages of learning a word, the new foreign language word has only one simple link to its L1 translation. The L1 word, however, has many associations, which is why productive recall is more difficult, as there are many competing paths the learner can choose from. To illustrate this point, consider an example for the foreign language Chinese with the L1 English. In the receptive direction this means when seeing the Chinese word 腿 [tuǐ] it links directly to its English translation *leg* but in the productive direction this means when seeing the English word *leg* and thinking of its Chinese translation 腿 [tuǐ], collocates, synonyms, opposites and other associations in the L1, which compete with each other, are co-activated and make the production of its Chinese counterpart more difficult because in productive recall there are more paths to choose from with the associations of the L1 likely being stronger. The fourth and last explanation includes the motivational factor, which is the reason for certain words or knowledge being known receptively but not used productively.

On the surface of it, for most learners knowing a word means knowing what a word sounds or looks like (i.e. its spoken or written form) but word form is of course not the only important aspect because the meaning of a word is also crucial. Therefore meaning and form need to be connected and the strength of this connection determines how easily the meaning can be retrieved when seeing or hearing the word form or how easily the word form can be retrieved to express meaning. This connection is strengthened by every successful retrieval of form or meaning, which is why it is important to give learners ample opportunities for such retrievals. Another factor exists that reinforces the connection of form and meaning: that is the case if the

sound or shape of a word form has an unambiguous even biunique connection to its meaning. (Nation, 2013)

A common problem arising when learning new words, especially high frequency ones, is that they have a lot of different meanings, which might be quite different from each other. If the different meaning of such a word can still be put in relation to each other, if only through sometimes complex processes of metaphorical extension, this is called *polysemy*. For example, the English word *bat* can have two quite different meanings: ‘flying mammal’ or ‘stick used in cricket/baseball’. Such words that have completely unrelated different meanings but the same (written and spoken) forms are called *homonyms*. There are also words that have the same written forms but different spoken forms, which are called *homographs*, such as the English word *row* in *row of people* (i.e. a line of people next to each other) and *make a row* (i.e. make a loud noise). In Chinese a large number of homographs (多音字 [duō yīnzì]) can be found as well. Consider the character 乐, for example: in the word 快乐 (= ‘happy’) it is pronounced kuàilè, while in the word 音乐 (= ‘music’) it is pronounced yīnyuè. Another example for a Chinese homograph is the character 还. This character can be pronounced hái (= ‘still’, ‘in addition’) or huán (= ‘to give back’, ‘to return’). The character 行 is a third example, as it can be pronounced xíng (= ‘to walk’) or háng (= ‘line’, ‘row’, ‘profession’). Additional to this, words that have the same spoken forms and different written forms also exist, which are called *homophones*. In English the words *blue* and *blew* would be an example for this type of word. (Nation, 2013:76) Homophones may be a problem for learners of CFL if they have not fully grasped the tonal system, as the phonemic shape of numerous Chinese words is identical due to the limited number of syllables. This might lead to incorrect meaning retrieval if the tone is ignored or misheard. Furthermore, an additional challenge is also posed by homophones who not only share the same phonemic shape but also share the same tone. An example of this would be the words 男 [nán] (= ‘male’), 难 [nán] (= ‘difficult’) and 南 [nán] (= ‘south’).

More on frameworks of word knowledge will be discussed in the section on depth of vocabulary knowledge.

Before moving on to the different dimensions of word knowledge, the burden of learning vocabulary in Chinese should be summarised here, as Chinese is often regarded as one of the most difficult languages to learn for learners with English as their L1. Learning Chinese vocabulary includes several difficulties, which will be discussed below. The first and most prominent feature that may cause difficulties to learners is the writing system, which was

discussed in section 3.4 *The Chinese writing system*. Due to the high number of Chinese characters achieving literacy may be rather hard. Another difficulty for learners with an Indo-European L1 is caused by the fact that Chinese is a tonal language. While tones are only used to express certain emphases or emotions in European languages, in Chinese tones distinguish meanings of words, which was shown in the introduction to the third chapter on the Chinese lexicon.⁷ However, it is not only difficult to retain the correct tones. An additional difficulty for non-native learners of CFL is using Chinese intonation at the same time. Other factors that also contribute to the learning burden of CFL are that there are different Romanization methods (e.g. Hànyǔ Pīnyīn) or that looking up words in a dictionary is not as easy as in European languages. (Moser, 1991)

4.2. Dimensions of word knowledge

It is important to understand that vocabulary knowledge is not a singular construct and learners may know different words to different degrees. Not only can vocabulary knowledge be divided into receptive and productive knowledge (see previous section) but vocabulary knowledge can also be divided into other dimensions, namely breadth and depth of vocabulary knowledge. The former considers learners' vocabulary size, while the latter considers how well learners know a word. The two dimensions will be discussed in more detail in the following sections.

4.2.1. Breadth of vocabulary knowledge

Scholars of L2 vocabulary learning agree that the size of learners' vocabulary is fundamental because it impacts them being able to use the target language. The comprehension of spoken and written discourse is crucially impacted by knowing the form and meaning of a certain number of words (e.g. Nation, 2006). Based on the distinction between receptive and productive knowledge, breadth of vocabulary knowledge has been distinguished into four categories: *form recall* (i.e. production of the L2 form of a given word meaning, e.g. by translating from L1 to L2), *meaning recall* (i.e. production of the meaning of a given L2 word, e.g. by providing a definition of this word), *form recognition* (i.e. recognition of the L2 form based on a given meaning, e.g. by selecting the meaning the best matching word to a given meaning) and *meaning recognition* (i.e. recognition of the meaning of a L2 word, e.g. by selecting the meaning of a L2 word).

⁷ As a reminder the examples used were the following syllables: 妈 [mā], 麻 [má], 马 [mǎ] and 骂 [mà]. These syllables are all pronounced [ma] but with different tones, which changes their meaning. The above examples' meanings are *mother*, *hemp*, *horse* and *to swear*, respectively.

These categories differ in the strength of the represented form-meaning connections, with form recall representing the strongest form-meaning connection, followed by meaning recall, form recognition and then meaning recognition. Consequently, being able to master each of these four categories is difficult to various degrees: recognition of the meaning of an L2 word is easier than its production, for example. (Laufer & Goldstein, 2004)

In order to be able to teach or learn L2 vocabulary in a systematic way, it is important to know how much vocabulary needs to be learned for L2 learners to be able to function in the target language. This is the reason why vocabulary-learning goals need to be set, which can be done in several ways. One of these ways is to look at how many words the target language contains. However, in chapter two it was discussed that it is not easy to define what a word is exactly, which consequently makes it difficult to define how many words a language has. So it should come as no surprise that different sources and researchers present different figures and estimates. If we consider the English language as an example, estimates on how many words it contains range from 54,000 word families to a million words. (Nation, 2013) Ultimately, however, such estimates of total vocabulary size are unhelpful for setting clear and achievable learning goals as it is unrealistic to learn all the words of a given language, let alone a second language. A second way of determining what size of vocabulary learners need is to look at what the size of native speakers' vocabulary looks like. Just as before with estimating the vocabulary size of a language, however, it is not as simple as it may sound, since the results will vary depending on the source used, the adopted definition of word and also testing methods. Educated English native speakers, for example, are said to have a vocabulary size of under 20.000 word families. (Nation, 2013; Schmitt & McCarthy, 1997) Now consider the vocabulary size of the Chinese language. The number of 'words' that the Chinese language contains is always estimated with characters not with words. The number of characters that is in use today is highly discussed and controversial. In the majority of cases the number of characters that is in use today is estimated to be ten to twenty thousand, which is remarkably similar to the estimates mentioned for English above. As mentioned in chapter three, Chinese children learn 2.500 to 2.800 characters in six years of school. (Guder-Manitius, 2009:45) According to Hobbs (2019) about 2.000 characters are necessary to read a Chinese newspaper. This can serve as an approximation for setting learning goals for CFL learners with regard to vocabulary size.

As with other foreign languages, there are language tests that were specifically developed for CFL, mostly in order to assess learners' language proficiency. These will be shortly discussed here.

4.2.2. Depth of vocabulary knowledge

The size of vocabulary, however, is not the only relevant factor in language learning, as it is not only important to determine how *many* words learners should know. Another relevant question is what constitutes word knowledge or how well learners know a word. The most evident aspect here is the connection between word form and word meaning, which shows that breadth and depth of vocabulary correlate with each other. Schmitt (2008) remarks that many learners and teachers believe that when learners know a word's written and phonetic form and meaning a word has been learned. There is no denying that these aspects are the most important and primary aspects of word knowledge. However, there is much more that learners need to know about words. (Liu, 2018) This is especially true for their productive use. In Nation's words "full mastery of a word requires more than just knowledge of its meaning and form." (2013:30-33) The different aspects of word knowledge are therefore central for teaching and learning vocabulary.

Unsurprisingly, a number of researchers have designed frameworks of these aspects of word knowledge. One of the earliest frameworks was proposed by Cronbach (1942). In his categorisation Cronbach (1942) distinguishes five types of behaviour that indicate learners' knowledge of a word. A limitation of this categorisation by Cronbach is that it focuses particularly on meaning and use but neglects word form (i.e. pronunciation and orthographic structure) and syntactic information.

Richards (1976:83) also proposed one of the earlier frameworks of word knowledge. The elaborations on his framework recognize the importance of: rapidly and massively expanding vocabulary in an L2 or foreign language context (i.e. assumptions one and eight), collocations (i.e. assumption two), awareness of differences between spoken and written language and different registers (i.e. assumption three), syntactic characteristics of words (i.e. assumption four), awareness of words' morphological features (i.e. assumption five), the networks between the words and their storage (i.e. assumption six) and, lastly, words' connotations in different contexts (i.e. assumption seven). An advantage of this framework compared to Cronbach's is that it includes additional important vocabulary features, such as frequency, register, collocational relationships and meaning associations between words. Richards (1976:88) concludes that awareness of the above aspects of word knowledge should feed into the teaching of vocabulary because our understanding of the nature of what is taught should always impact its teaching.

Other existing definitions of word knowledge and its constituents have been summarised by Schmitt and McCarthy (1997:141). Their summary includes the following six points:

1. Spoken and written word form (i.e. pronunciation and spelling)
2. Word structure (i.e. basic word form and its derivations and inflections)
3. Syntactic pattern of the word (i.e. in a phrase and sentence)
4. Referential, affective and pragmatic word meaning.
5. Lexical relations of the word with other words (e.g. synonyms, antonyms, hyponyms, etc.)
6. Common collocation of the word

In order to gain a more practical understanding of what the aspects of knowing a word entail, a look at Nation's (2013) framework is helpful. In this framework Nation divides the aspects of word knowledge into three main categories: word form, word meaning and word use. Each of these main categories has three subcategories, which are each also divided into receptive and productive vocabulary knowledge catering for the distinct abilities that each of them requires. Table 2 shows in detail what is included in each subcategory of this framework, which Nation formulated as questions.

Note: *R* = receptive knowledge, *P* = productive knowledge

FORM	spoken	R	What does the word sound like?
		P	How is it pronounced?
	written	R	What does the word look like?
		P	How is the word written and spelled?
MEANING	word parts	R	What parts are recognisable in this word?
		P	What word parts are needed to express the meaning?
	form and meaning	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	concept and referents	R	What is included in the concept?
		P	What items can the concept refer to?
	associations	R	What other words does this make us think of?
		P	What other words could we use instead of this one?
USE	grammatical functions	R	In what patterns does the word occur?
		P	In what patterns must we use the word?
	collocations	R	What words or types of words occur with this one?
		P	What words or types of words must we use with this one?
	constraints on use	R	Where, when and how often would we expect to meet this word?
		P	Where, when and how often can we use this word?

Table 2. What is involved in knowing a word? (Nation, 2013:49)

In order to better understand this framework consider the following example of what is involved in knowing the word ‘underdeveloped’ (Nation, 2013:48-50) split into receptive and productive knowledge and use below. According to Nation (2013) knowing the word ‘underdeveloped’ involves the following abilities:

Receptive knowledge and use

- recognising the word form when reading or hearing it (i.e. being familiar with the word form)
- recognising its word parts (i.e. under-, -develop-, -ed)
- recognising the particular meaning signalled by the word
- knowing the word's meaning in a particular context
- knowing the concept behind the word
- knowing related words (e.g. 'overdeveloped' or 'backward')
- recognising the correct use of the word
- recognising typical collocations (e.g. 'territories' or 'areas')
- knowing it is not an uncommon word

Productive knowledge and use

- saying the word (with correct pronunciation and stress)
- writing the word (with correct spelling)
- constructing the word (using the right word parts in their appropriate form)
- producing the word to express its meaning (in different contexts)
- producing synonyms and antonyms of the word
- using the word correctly in a sentence
- producing words commonly occurring with it (i.e. collocations)
- deciding whether to use the word or not depending on the context

Nation's taxonomy of elements of word knowledge is the most comprehensive and most systematic one to date and it shows that knowing a word includes quite a large number of aspects and therefore is a highly complex matter, which goes beyond the link between form and meaning. This means that in general, but in foreign language learning and testing contexts specifically, it is rather unrealistic to tackle all the aspects of word knowledge at once. This has implications for the empirical study undertaken in this thesis, as only a limited number of aspects of CFL learners' word knowledge can be operationalized in a single study.

All frameworks discussed in this section share the assumption that knowing a word does not mean knowing everything about a word. There are different degrees to which one can know a word. However, the more information one knows about a word, the better one is able to understand it. This emphasizes again the fact that vocabulary breadth (i.e. number of words one knows) and vocabulary depth (i.e. to what extent one knows each word) should be two different but related concepts.

4.3. Types of vocabulary learning

So far it has been established that not only do learners have to acquire a large number of words but that they also need to learn a large amount of information about each individual word. Unsurprisingly, therefore, a large number of researchers have examined the best ways to learn vocabulary. Literature on L2 acquisition highlights two main types of learning: incidental and intentional learning. The former occurs (mainly) unintentionally. This means that learners may learn new words from context while doing other tasks, such as reading a book or participating in a conversation. L1 vocabulary acquisition predominantly happens in this manner. Intentional learning on the other hand refers to the deliberate learning of vocabulary. This can happen by working with word cards or by completing vocabulary activities in a textbook, for example. Multiple studies have tested the effectiveness of these two types of vocabulary learning and while some scholars advocate one method over the other, there is widespread agreement that each of them serves its purpose. (Nation, 2013) In the following both benefits and limitations of incidental and intentional learning will be discussed.

Intentional learning is one of the most researched types of learning and is mostly associated with studying from word cards or vocabulary lists, as mentioned above. It is, however, viewed critically by some teachers and researchers as it is seen as a de-contextualised way of learning and is not in agreement with the communicative approach to language learning that is popular nowadays. Especially the fact that word cards de-contextualise vocabulary is the main source of criticism of researchers. Nation (2013:438) states that this includes the following two points of criticisms: first, learning from word cards is not good for remembering. And secondly, learning from word cards does not help with the use of the word. Enough evidence exists, however, to counter the first statement. A review of studies by Nation shows that a great number of words can be learned in a short amount of time when presented without a context and that deliberate learning of vocabulary also leads to long term retention. According to Nation the second point of criticism is only partially true because intentional learning from word cards can help acquire three different aspects of word knowledge: written form, meaning and concept. Furthermore, word cards can also provide the learner with at least partial information of aspects such as grammatical features and collocations. (Nation, 2013; Nation & Webb, 2011)

In contrast to the criticisms mentioned above, one benefit of learning from word cards certainly is that it is focused, efficient and certain. If we consider Table 2 (*What is involved in knowing a word?*) again, Nation (2013:440) states “learning from word cards helps with learning the written form of the word, learning the concept of the word and making the connection between the form

and the meaning.” Learning from word cards, therefore covers three of the nine aspects involved in knowing a word.

The incidental approach to learning vocabulary is mainly associated with extensive reading. This type of learning has a number of benefits (see Nation, 2013; Schmitt, 2008). First, extensive reading gives learners opportunities to encounter new words in a variety of contexts. This way, learners can gradually expand their knowledge of the different aspects of the words. Secondly, meeting the same words repeatedly helps the learners to commit them to their long-term memory. Studies show that reading extensively in a classroom environment leads to the expansion of the learners’ vocabulary. Another way of learning vocabulary in a second or foreign language incidentally is watching films or series with subtitles. However, studies have shown (e.g. Birulés-Muntané & Soro-Faracao, 2016) that watching subtitled films has significant effect on improving learners’ listening comprehension, while the results were not really conclusive for vocabulary learning and showed only modest improvements. Other studies (e.g. Vulchanova, 2015) have shown that also vocabulary learning has increased after watching subtitled films. Thus, showing that researchers have mixed opinions on whether watching subtitled films has an effect on improving learners’ vocabulary.

While reading is indeed a valuable asset in expanding vocabulary knowledge, this method also has certain limitations. The main concept of extensive reading that is criticised by research is that new words are learned by guessing their meaning from context. Research suggests that learning can only take place if an unfamiliar word is recognised and its meaning is then inferred correctly, which is not always the case (Laufer, 2003:568). Some studies show that few new words are learned by reading extensively and that it rather helps to deepen the knowledge of already (partially) known words by repeating them multiple times, which improves cognitive entrenchment. Guessing from context successfully is further related to lexical coverage. It appears to be the case that 95-98% of the running words of a text need to be already known by the learners for them to be able to infer the meaning of unfamiliar words in a text. (Nation, 2013; Nation & Webb, 2011)

Nation (2013) points out that another limiting factor of extensive reading is the fact that it is very time consuming. The reason for this is that incidental learning and accumulating word knowledge through it can only take place through repeated encounters, which in return can only be ensured by enormous amounts of reading.

In view of this conflicting evidence it seems impossible to say that one type of learning is more effective than the other. In fact many researchers share the opinion that both types of learning play an important role in vocabulary learning and should complement each other (e.g. Nation, 2013). It needs to be said, though, that most studies dealing with this question focus on vocabulary learning in English as a foreign language (EFL) and can be summarised as follows: While word cards allow the learning of a great number of words in quite a short period of time, only partial word knowledge can be acquired via this method. Which is why other aspects of word knowledge are better learned by the means of incidental learning from context. Incidental learning also strengthens the learners' knowledge through repeated encounters with words, which cannot be done by intentional learning, as it is not possible to recycle words adequately. (Nation, 2013; Schmitt, 2008)

4.4. Important factors for learning vocabulary

Identifying the factors that lead to variations of foreign language acquisition processes has become more and more important since the 1970s starting with work on the “good language learner”. This research area is based on the assumption that “learners differ in how successfully they adapt to, and profit from instruction” (Robinson, 2002:9). Thus, this research area concerns itself with searching for explanations for the individual variations regarding learning success. (Burwitz et al., 2016)

This goal has led to different classification systems of said factors. Skehan (1989) is of the opinion that the following four factors impact language learning: *aptitude*, *motivation*, *learner strategies* and *learner styles*. While Oxford and Ehrmann (1993) also include aptitude, motivation and learning styles into their classification they also include six other factors. These are *anxiety*, *self-esteem*, *tolerance of ambiguity*, *risk-taking*, *age* and *gender*. Some researchers (e.g. Riemer, 1997), however, are of the opinion that the factors influencing language learning have to be classified more clearly. Thus, consider Riemer's (1997) classification system as an example. In this classification the factors are first distinguished into *individual* and *universal factors*. Each of these can be further distinguished into sub-categories. First, individual factors can be distinguished into *linguistic* and *non-linguistic factors*. Linguistic factors concern the learners' linguistic resources (i.e. all languages they have learned and their language learning experiences). According to Burwitz et al. (2016:243), studies show that if a learner has learned one or more foreign languages, they possess a higher language (learning) awareness. The non-linguistic factors can yet again be subdivided: into *learner-endogenous* and *learner-exogenous factors*. Learner-endogenous factors include *affective* (e.g. motivation, anxiety and attitude),

cognitive (e.g. language learning aptitude, learning styles and strategies) and *biological factors* (e.g. age and gender). The learner-exogenous factors include the learner's environment. On the one hand this includes the teaching environment (e.g. teacher, methods or materials) but on the other hand this also includes the learner's social environment.

Six of the factors that influence language learning that were mentioned above will be discussed in the following: age, gender, language learning aptitude, affective factors and social factors. First, consider the factor *age*: this can mean chronological age-related neurobiological maturation processes, as well as age-related changes in cognitive-affective and socio-cultural condition of L2 acquisition. Together with the factors of language learning aptitude, affective and social factors, the factor of age is central in explaining individual difference of L2 acquisition. Research shows that the age one starts learning a language impacts the competency that can be achieved: the later an intensive language acquisition process starts, the less likely it is to reach (almost) native language competency. However, the question of how to explain the fact that L2 learners who started learning at a younger age reach a higher language level is controversial. The central hypothesis here is that critical or sensitive phases exist for language learning, i.e. certain phases are better for language acquisition than other from a neurobiological standpoint. Researchers who share this standpoint generally are of the opinion that such a sensitive/critical phase is somewhere between the ages of three and fifteen years. Reasons for presuming a sensible/critical phase are the age-related decline of the brain's plasticity, age-related limitations in accessing the innate language acquisition mechanism or development-related decline of the ability to implicitly process and learn language. There is also a large number of researchers disagreeing with this hypothesis. They connect the age-related differences to other factors as well (e.g. motivation, influence of L1 or language aptitude). (Burwitz et al., 2016)

The second factor in language learning that will be discussed here is the role of *gender*, which is still considered unclear to this day. Nevertheless, a number of studies have indeed dealt with the topic of 'gender' from different angles. Burwitz et al. (2016:254-255) divide these studies into three different areas: sex/gender and learning success, motivation and attitudes and sex/gender and interaction. First, the most popular assumption (at least until the late 1990s) concerning *sex/gender and learning success* is that of a 'female superiority' in foreign language learning. Since the end of the 1990s there is a tendency to a more differentiated approach in this area, as it is not just assumed that girls are 'better' or 'more talented'. Studies actually show that there is no difference between genders when it comes to language learning. Second, a similar picture

can be found in the area of *motivation and attitudes*. Although female learners choose foreign language courses and studies more often than male learners, research has not been successful in finding gender-specific motivational profiles. Similar to the area of learning success, systematic differences between genders cannot be proved. Third, consider the studies on *gender and interaction*, which try to conceptualise gender as a 'discursively created category', as the biological sex cannot be seen as the crucial factor for the social and cultural development of genders. In the area of foreign language learning and teaching this has also led to new research topics.

Language learning aptitude is the third factor to be considered in this context. In the beginning research concentrated on finding testing instruments for language learning aptitude but not on finding a definition of this notion, which is a pivotal deficit of this research area. Today, still no consensus could be found which components make up language learning aptitude. Unity can only be found in that it can be defined as a "mehrdimensionales, komplexes, verschiedene kognitive Fähigkeiten umfassendes Konstrukt" (= 'multi-dimensional, complex construct that includes different cognitive abilities') (Burwitz et al., 2016:258). Thus, the question arises whether research has found some components that can be said to play a role in language learning aptitude. Research indeed shows that working memory is an essential component of language learning aptitude. According to Burwitz et al. (2016) two models are of relevance. First that of Skehan (2002) has conceptualized a "*Processing Stage Model of Aptitude*", which has four macro-phases: noticing, patterning, controlling and lexicalising. Each of these is further divided into sub-stages of acquisition, making a total of nine stages. Learning aptitude components (e.g. segmenting, working memory, grammatical sensitivity or chunking) are then assigned to each of these stages. So-called *aptitude complexes* have been conceptualised by Robinson (2001), which arise from the interaction between individual cognitive abilities and the specific learning situation.

Another factor that has impact on language learning are *learning styles*. This notion, as others mentioned before, does not have a consistent definition, which hinders the clear distinction between learning styles and other factors. Nowadays the notion learning style is often seen as a superordinate concept, which includes cognitive style, sensory preferences and personality types. However, some researchers (e.g. Dörnyei & Skehan, 2003) see cognitive style as a way of perceiving and categorising information preferred by an individual across different situations, while learning style describes a preference of dealing with information in a certain learning context or situation. According to Burwitz et al. (2016) research has identified a variety of

different dimensions of learning styles. However, it is important to note that specific learning styles do not yield better or worse competence: different learners can be equally successful with different learning styles.

Crucial to anything humans do are *affective factors*, such as attitudes, motivation and anxiety or fear. Because learning a foreign language takes years, these factors can contribute to explaining the differences in attainment. *Attitudes* of learners towards the L2 or foreign language are often researched in connection to motivation. However, early studies have shown that a positive attitude towards the language does not necessarily mean more successful learning. *Motivation* is the most researched affective factor in the context of language learning. This includes research on the reason why learners learn a certain language, as well as on motivational processes while learning the language. Earlier research concentrated on whether attitudes towards and interest in the language and the culture connected to it had impact on the learning of said language (i.e. socio-educational model). Today, the connection between language learning, motivation and identity development of learners is in the focus of research. The third affective factor is *language-specific anxiety or fear*, which has also been researched intensively. However, studies have shown little influence of this factor on successful foreign language learning. This notion includes anxiety/fear of speaking, understanding, social anxiety and anxiety/fear in test situations. (Burwitz et al., 2016) Most recently a new research perspective of *foreign language enjoyment* has been added. (De Waele et al., 2017)

Social factors refer to external influences on learners and their learning processes. The questions of whether such factors promote or impair the language learning process and whether such factors play an important role at all has been the focus of research for a long time. Burwitz et al. (2016:271) state that the distinction between *external* and *internal learner factors* is still common. External factors are those aspects that cannot be associated directly with the individual, cognitive processing, for example social class or ethnicity.

Lastly, in context with foreign language learning it is also important to consider learners' stays abroad. The idea of learning a foreign language 'incidentally' while staying abroad is a rather popular one. The popularity of stays abroad lies in the fact that improving language abilities and gaining cultural experiences can be combined. However, a multitude of factors influence language learning, thus a stay abroad on its own cannot be a guarantee for successful language learning. Due to the highly complex and interdisciplinary nature of this area, a lacking amount of research has been done. Coleman (2010), however, has created a study abroad taxonomy containing 20 parameters (e.g. age of participants, type of stay, accommodation,

preparation, duration and content orientation). Concerning improvement of language competency, stays of several months of students have been researched. These studies show that language competency in total improves during such a stay, however, the improvement of partial language skills depend on various factors (i.e. individual, institutional or research-methodological factors). Research on school exchange programmes primarily focuses on intercultural learning processes. These studies show that a preparation, support and a follow-up of these exchange encounters are important. Due to the various different types of study abroad experiences a large amount of research is still needed in this area. Due to the role of individual factors, the generalizability of results is difficult. Thus, they can be complemented by case studies in order to better define the learning, working and research area of stays abroad. (Burwitz et al., 2016)

4.5. Learning vocabulary in CFL

After discussing vocabulary learning in general, this section will focus on the learning of vocabulary in a CFL context. In chapter three the lexicon of the Chinese language was discussed, including words, collocations and the writing system. Now the focus will turn to what this means for learners of CFL and how they learn vocabulary, especially focussing on word recognition, as this study implements a receptive written vocabulary test. Additionally, the majority of studies focus in some way on Chinese characters and how they are recognised by CFL learners.

Learning vocabulary in any foreign language can have its challenges but Chinese poses unique challenges to learners with an Indo-European language background due to its language features, which were discussed in chapter three. First due to the features of the writing system, word recognition is difficult – especially for beginners. This includes recalling word meaning because according to Perfetti et al. (1992) it is believed that the access to word meaning is mediated through sound and in Chinese sound information is not readily available from the written form. Although the semantic components of characters might help in providing hints about word meaning, these represent meaning boundaries rather than accurate meaning information. Sometimes the connection of a character's meaning and its semantic component is so loose that the component is not helpful in word recognition at all. In the following, studies on word recognition and their results and implications for learning CFL will be discussed in more detail.

Everson's (1998) study on word recognition focuses on how learners derive pronunciation and lexical meaning from written/printed characters. Additionally, Everson tries to assess the relationship between correct pronunciation and correct identification of words. The participants of this study are twenty American CFL learners at beginner's level, who are all non-heritage learners of Chinese. Japanese and Korean speakers were also excluded due to the familiarity of writing systems. At the time the study was conducted the participants had formally studied approximately 500 characters. The study consisted of the following two tasks: first, the participants were asked to pronounce Chinese words aloud as quickly and accurately as possible. In the next step they were asked to write the English meaning next to each word. The results of this study show that pronouncing and identifying words is rather challenging for beginning learners of CFL and that there is a high correlation between pronunciation and identification of Chinese words. The participants of this study had a mean probability of 90.7% of correctly identifying the meaning of a word, when they are able to correctly pronounce it, as well as a mean probability of 91.4% of pronouncing a word correctly, when they identified it correctly. These results demonstrate that a strong relationship between knowing a word's meaning and its pronunciation, at least at this level of language proficiency, which is consistent with other studies. Vocalisation of words is also a strategy to better isolate the words within a written text, as there are no extra spaces between characters.

This connection between speech and meaning has also been shown in two studies on reading strategies. The first study was conducted by Everson and Ke (1997) on intermediate and advanced learners of CFL. This study showed that a lot of time was spent on decoding unknown characters. Additionally, the participants at the intermediate proficiency level had difficulties isolating meaningful word units but the advanced students did not. Lee-Thompson (2008) conducted the second study on reading strategies and on learners' difficulties while reading. The study's participants had an intermediate proficiency level, were native speakers of English and had learned Chinese for at least two years. Lee-Thompson found that there were four types of difficulties learners had when reading Chinese texts: vocabulary, orthography, grammar and background knowledge. Consider the difficulties concerning vocabulary, as this is also relevant to the context of this thesis. The results of this study show that most of the participants' focus was put on understanding individual words and not the meaning of larger portions of text, which was probably due to the high quantity of unknown words. This shows the crucial role of vocabulary knowledge in a CFL context. 50% of the study's participants also had difficulties recognizing and remembering Chinese characters and words. This is (mainly) due to the orthographic difference of Chinese and English. Additionally, this study also found learners

attempted to pronounce a word while trying to retrieve its meaning, which is further evidence that phonological access seems to be an essential component of word recognition in Chinese.

Another aspect shown to facilitate word recognition is radical knowledge. This was the case in Shen's (2000) and Williams (2013) studies, which also showed that especially the knowledge of semantic component or radicals is helpful for recognising meaning in morphologically transparent characters. Additionally, reading time and accuracy were facilitated as well when the semantic component clearly indicated the semantic category of the character. Correspondingly, when the meanings of the radical and the character are non-congruent, reading time and accuracy are impaired. This goes to show that clearly, character acquisition is an important part of vocabulary learning in a CFL context and that knowing and using Chinese radical knowledge can facilitate character acquisition. Consider again the two types of learning (i.e. intentional and incidental learning) discussed in section 4.3. *Types of vocabulary learning*, it can be said that Chinese characters are most commonly learned by intentional learning, as it is central to learn their meaning and pronunciation. Thus, characters and words are often learned as isolated units. Contextual learning often happens by means of incidental learning without using characters, as a lot of CFL learners are able to use more words when speaking than they are able to read or write. In consequence, when reading, words are often recognised by their phrasal context.

Another study in this area is the one conducted by Shen and Ke (2007) on radical awareness⁸ and word acquisition. This study's participants also were American, non-native learners of Chinese from beginning to advanced levels; heritage learners and learners with other Asian language backgrounds were again excluded. The study included 35 participants for each proficiency level. Shen and Ke include three levels of radical knowledge into their study: radical perception (i.e. knowledge of how to decompose characters into radicals), radical knowledge (i.e. knowledge of the semantic radicals' sound, shape and meaning) and application of radical knowledge (i.e. knowledge of how to use semantic radicals in learning compound characters). The results of the study show that the trend of acquisition of semantic radical knowledge is linear; meaning that the participants' radical knowledge increased as the learning level advanced. The radical perception skill also shows a linear development trend until a certain point (which is about one year of learning), with no significant development after said point. The

⁸ *Radical awareness* is "the functional understanding of the role of radicals in forming Chinese characters and the ability to use this knowledge consciously in learning characters. (Shen & Ke, 2007:100)

use of radical knowledge in learning new characters shows two periods with linear trends: from one month to one year and then again from the end of the second year to the end of the third year of learning. In between the first and second year there is little progress in this skill. Shen and Ke also considered the question of whether radical knowledge plays a role in the development of radical perception and using said knowledge when learning new characters. Results show a linear trend between both radical perception and radical knowledge application and radical knowledge. The skills of decomposing characters into radical units and reproducing characters out of radical units emerges at the very beginning (participants show these skills also after only one month of learning) and develop rapidly during the first year of learning Chinese. By the end of the first year the accuracy rate of these skills lies at about 73%. According to the authors this can be due to two reasons. One is that the learners are adults and therefore cognitively mature and able to analyse radicals within a character rather quickly. The second is that the participants have experience with the graphic structure of compound characters because they have been taught how characters are constructed. However, at the beginning learners still tend to recognise characters as a whole rather than recognising their parts, as they have insufficient orthographic knowledge at that stage. Developing radical perception skills can reduce the burden of character acquisition and memorisation because this helps our brain to systematically process characters. This skill develops rather slowly, though, as the study shows that after three years of study the participants had a 70.95% mean accuracy rate for 40 high-frequency radicals. The reason for this may be that learners need to form new cognitive structures to assimilate Chinese radicals.

Second, the concept of a word and how words are constructed in Chinese also impacts CFL learners. This is because Chinese orthography does not segment written texts into words, which is the case for languages such as English or German. Chinese, however, segments texts into characters, which represent morphemes rather than words. Consequently, word boundaries are not directly perceivable, which adds difficulty in determining what constitutes a word. Not only CFL learners have difficulty with separating word boundaries, also Chinese native speakers do not always agree on how a written text is segmented into words. CFL learners generally seem to rely on one main strategy of word separation: translation into L1 (which in most studies is English). Native speakers on the other hand apply a range of different strategies. According to Bassetti (2005) this is also the reason why native speakers seem to have some differences of opinion concerning word boundaries. Furthermore, CFL learners in Bassetti's study tended to identify shorter words than native speakers. This was also attributed to the influence of the L1 writing system (i.e. English). However, there is also evidence that CFL learners begin to

develop word awareness similar to the Chinese native speakers. In another study Bassetti (2009) also asked learners of CFL and Chinese native speakers to determine matches between a picture and a sentence written in Pīnyīn or Chinese characters with either inter-word or inter-syllable spacing. This study found that while inter-word spacing tended to facilitate Pīnyīn reading among CFL learners but not native speakers, it had no effect on reading in Chinese characters. (Packard, 2000; Perfetti et al., 1992)

The following study by Shen (2008) is also interesting in the context of vocabulary learning. The focus of this study lies on word decision strategies for unknown words of 40 beginning and advanced American CFL learners. The study included a character knowledge worksheet to measure the participants' character knowledge and a word decision task. The target words were unfamiliar, however, the constituent characters that make up the target words were familiar. To ensure readability of the texts the percentage of unknown words was below 10%. In her analysis Shen identified nine types of strategies from the student's responses. These included guessing, matching target constituent characters to existing ones in the mental lexicon or combining semantic information of constituent characters of the target items. The frequency of strategy use was consistent in both groups, with matching target constituent characters to existing ones in their mental lexicon being the most frequently used by both groups of learners (30% and 29%, respectively). Activating relevant linguistic knowledge (i.e. phonology, orthography, semantics and grammar) for words or sentences was involved in making word decisions. However, word decision is a difficult task in a CFL context, even for advanced learners. Results show that the word decision accuracy rate of the beginning learners was 50% and for the advanced learners 54%. This shows a rather slow progress in developing the skills necessary for word decision. On the one hand this is due to the complexity of Chinese orthography, as linguistic factors (e.g. morphological structure of words or parts of speech) affect word decision. On the other hand psycholinguistic constraints in lexical development and representation also affect word decision skills. This includes limited input of lexical knowledge and heavy dependence on lexical knowledge and learning behaviours of the learners' L1. This shows that correct word decisions depend on the students' confidence, as well as their ability to draw on their linguistic knowledge of the Chinese language and their experience with Chinese words. Therefore, it is not only necessary to introduce students to a large quantity of words but also helping them to understand that a character can have many different meanings when it is grouped with other characters. Additionally, grammar knowledge (e.g. parts of speech, sentence components or word formation) is also important for reading.

Earlier in this chapter vocabulary breadth and depth in general were discussed, my focus will now turn on the CFL studies in this area, which are few in number. First, Lü's (2010) study looked at word associations by the means of a word association task, which will be further discussed in the next chapter on the assessment of vocabulary. These data were then used to examine how CSL learners acquire different dimensions of vocabulary (i.e. vocabulary depth) and also how this differs across different levels of language proficiency. At a higher proficiency level, the number of syntagmatic and paradigmatic responses increased but did not shift from the former to the latter (as was assumed by the author of the study). Familiar words elicited more syntagmatic and paradigmatic responses than unfamiliar words, which elicited more unrelated responses. Although this study provides us with interesting and useful insights into how CSL learners develop depth of vocabulary knowledge, its validity is impaired by a methodological problem because the participants were asked to produce the associates on their own in a very limited time frame (i.e. ten seconds).

Second, the study by Chen (2016) is one more study focussing on the depth of vocabulary knowledge of advanced CFL learners. As this study is the model study on which my own study is based it will be discussed in detail in chapter six on research methods.

Third, Zhang and Yang (2016) explored the depth of vocabulary of CFL learners and its contribution to their reading comprehension. Heritage learners were again excluded from the study, however, the participants had different L1s (e.g. English, Russian or Korean). They were studying in China at the time of the study (86% of the participants had studied in China for one year). Additionally, 76% of the participants had learned Chinese in their home country for about two years or more prior to studying in China. This study includes a variety of measures (e.g. a background survey, a Word Associates Test (on synonyms and collocations), an orthographic processing task or a picture selection task). This study's results show that word frequency did not impact the learners' test performance. The fact that word frequency did not impact the learners' test performance may be due to the fact that all words in the WAT were chosen from the most frequent 2.000 words for entry-level Chinese learners. Word class and type of association relationships also had no significant impact on the WAT results, as the analysis showed that no significant difference could be found between adjectives and verbs and paradigmatic and syntagmatic associations, respectively. Depth of vocabulary knowledge on the other hand was a significant predictor of reading comprehension test performance. However, the contribution of vocabulary depth also depended on the types of texts the participants read and what comprehension skills were assessed. Vocabulary depth showed to be an important

predictor for inference, which was tested by short passage comprehension tasks. For literal comprehension, which was tested with a long passage comprehension task with questions, both vocabulary size and vocabulary depth were significant predictors (but results showed that vocabulary size was a more important predictor in this case). Based on these results implications for teaching CFL are drawn as well. Zhang and Yang (2016) are of the opinion that it is indeed important for the learners to develop an understanding of the basic meanings of a large number of words, however it is as important to build a network of the words' meanings and how they relate to each other. However, textbook analyses show that attention is mostly put on forms and meanings. Although it is important to note that these analyses are not of Chinese textbooks, as it is unknown which aspects of vocabulary these textbooks focus on, according to Zhang and Yang. Reports do show, however, that CFL learners often lack "fine-grained understandings of word meanings, such as nuanced meanings of synonyms and their distinct collocation patterns" (Zhang & Yang, 2016:712).

5. Assessing vocabulary knowledge

When testing vocabulary the same criteria apply when testing other areas of language knowledge or language use: reliability, validity, practicality and washback, which should also be considered for the design and the evaluation of a vocabulary test. For designing a test what needs to be considered is the purpose of the test, as well as what knowledge it tries to measure and which condition it will be used under. As mentioned, different purposes of tests exist for language or vocabulary knowledge and use. We shall discuss the purposes of language tests here shortly before going into different tests and what they are used for. First, if learners experience difficulties, language tests can be used to find out what these difficulties are and then something can be done about it; this type of test is called diagnostic test. Secondly, more or less every language learner has some sort of experience with the next purpose of tests, which is a placement test. It is used to find out at which level of a language course learners should be placed. But language tests can also be used to see whether certain words have been learned. These tests can be divided into short-term achievement tests, which look at whether a recently studied group of words has been learned, and long-term achievement tests, which are used to see whether the words of a particular course have been taught successfully. The last purpose is to test language proficiency to see how many words learners know. Thus, the type

and purpose of a test determine which vocabulary is selected for inclusion in the test items. (Nation, 2013)

The literature on vocabulary testing emphasises that a good vocabulary test contains plenty of

Form	spoken	R	Can the learner recognise the spoken form of the word?
<i>Note on column 3: R = receptive knowledge, P = productive knowledge</i>			
		P	Can the learner pronounce the word correctly?

items, with a minimum of thirty items for the test to be reliable (cf. Nation, 2013). Furthermore it needs to be ascertained that test items require the test takers to use the kind of knowledge one wants to test. A general consideration of possible test items shows that they can differ in many ways because some tests give choices, some use the L1, some tests put the word into a sentence context and some require the test taker to use the word. Also the aspects of vocabulary knowledge which test items can focus on can be different. Some focus on meaning, some on word form, while others focus on grammar, collocations or associations, which is shown in more detail in Table 3 on the next page. (Nation, 2013)

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

On the basis of the table above different types of test items can be developed, as Table 3 is based on receptive and productive aspects of vocabulary knowledge, the following test items created by Nation (2013:551-552) are as well.

- | | |
|--|---|
| <ul style="list-style-type: none"> • Spoken form <ul style="list-style-type: none"> (R) Word / sentence dictation or choosing the L1 translation after hearing the word (P) Reading aloud or (cued) oral recall • Written form <ul style="list-style-type: none"> (R) Saying written word or saying regularly spelled nonsense words (P) Word / sentence dictation • Word parts <ul style="list-style-type: none"> (R) Breaking words into parts or choose / provide the meanings of the parts (P) Providing an affixed form of a known word • Form and meaning <ul style="list-style-type: none"> (R) Translating words into L1 or choosing the right picture (P) Translating words into L2 • Concept and referents <ul style="list-style-type: none"> (R) Translating underlined words into L1 (P) Choosing the words to translate the given L1 word | <ul style="list-style-type: none"> • Associations <ul style="list-style-type: none"> (R) Choosing words that you associate with the given word (P) Adding to a list of associated words • Grammatical functions <ul style="list-style-type: none"> (R) Is this sentence correct? (P) Use a given word in a sentence • Collocations <ul style="list-style-type: none"> (R) Is this sentence correct? (P) Produce collocations to go with a given word • Constraints <ul style="list-style-type: none"> (R) Which of these words represent UK use? (P) What is the formal word for X? |
|--|---|

From the above list of test item types it becomes clear how varied tests for vocabulary knowledge can be. Therefore the researcher or test developer needs to consider carefully what aspect of vocabulary knowledge should be assessed by the test and choose the type of test item accordingly. We will now take this study as an example. As mentioned in the introduction, this study looks at collocations in the Chinese language; the methodology behind this study will be discussed in more detail in chapter six. If we now look at the right column of the list above,

we can see what types of test items are listed under collocations: to test receptive knowledge of collocations the test taker has to decide whether a sentence or collocation is correct. To test productive knowledge of collocations the test taker has to produce collocations that go with a given target word. When considering this and Table 3, we can see the present study tests for receptive knowledge of collocations, as one research question is how well students of Chinese Studies can identify Chinese collocations. More details on what test items were chosen for this study will follow in chapter six.

A criterion that cuts across the receptive and productive categories is the question of how *sensitive* a test is. In the context of vocabulary tests the term *sensitive* refers to whether or not a test allows for partial knowledge to be used to answer the test items and gives credit for it. For example, a multiple-choice test is more sensitive than a translation test. The reason for this is that a multiple-choice test allows for meaning recognition, while translation tasks fall into the meaning recall category.

Apart from being sensitive a good vocabulary test has to fulfil two important criteria that have not been discussed before: *reliability* and *validity*. Generally speaking, the term reliability describes whether a test measures consistently and accurately – including whether a test is stable over time. Another important factor concerning tests is their validity, which as Milton (2009:18) puts it means “whether a test measures what it is supposed to measure and not something else.” Different kinds of validity can be distinguished, which concern different aspects of a test’s validity. On the one hand we have *content validity* that, as suggested by the term itself, concerns the content of a test and whether or not it is appropriate for measuring what it is supposed to. On the other hand, whether or not a test measures the construct or skill it is supposed to is covered by the term of *construct validity*, which is closely related to content validity. This is rather tricky because language knowledge is not directly accessible. Therefore, language production has to be used to infer measurements of language knowledge, which can be problematic as it may involve other knowledge and abilities than only vocabulary knowledge itself. The third type of validity that should not be neglected is *face validity*, which concerns the credibility of a test (and what it is supposed to measure) in the eyes of the test takers. During the development of a test designers have to make sure that it turns out to be both reliable and valid. This is achieved by looking at the two main concerns: selection of words to be tested and method for checking the test-takers’ knowledge of the chosen words. A certain consensus has been reached among vocabulary testing experts (Milton, 2009; Nation, 2013) on the issue of word selection by using word frequency data to decide on which words should be included in a

test. Naturally, preference is given to testing the most frequent vocabulary of a language. A general consensus is not possible or even useful for the second issue of the implied method, which I have already discussed in connection with Table 3. The test of the current study also shows problems concerning the frequency of the included words, which will be discussed in the sixth chapter on research methods.

5.1. Assessing vocabulary breadth

With vocabulary breadth being the most basic dimension of lexical competence, it is no surprise that a number of tests exist in this area. The majority of these tests focus on receptive knowledge, as it normally precedes productive knowledge, in both L1 and L2. The reason for this is that certain contextual and textual clues and information are only available during reading and listening but not in speaking or writing. A good vocabulary test makes use of word frequency, as learners are more likely to know the most frequent words in a language. By doing so estimates about the words learners have most likely encountered and therefore about their knowledge of said words can be made. In accordance, estimates about the usefulness of learners' vocabulary for comprehension and communication can be made as well. The aspect most vocabulary breadth tests examine is single word meaning. To achieve this goal tasks such as matching, multiple choice or translation are employed.

The tests we will focus on in this section are tests of written receptive vocabulary size, which includes the following three most commonly used tests: the Vocabulary Levels Test, the Vocabulary Size Test and the Yes/No Test. But many more tests for vocabulary breadth – both receptive and productive – exist, such as tests for aural receptive and written productive vocabulary size. (Beglar & Nation, 2014; Milton, 2009) As very many of these tests have been developed for the English language only, I will discuss only the three most common types as test construction of other languages is likely to be modelled on them.

5.1.1. Yes/No Test

Originally the *Yes/No test* (Meara, 1989) was conceptualised as a placement test, which is why early research focused on developing a computerised version of this test format. The outcome of this research was the *Eurocentres Vocabulary Size Test* (Meara & Jones, 1990), which contains 150 words to estimate learners' knowledge of the 10,000 most frequent lemmas of the English language.

The general format of this test is a checklist containing a combination of real and non-words. In this checklist the test takers mark the words they know (e.g. by putting a tick next to the words). The reason why non-words are included is to adjust for test takers' guessing. As can be seen in Illustration 1 below, this test format is a rather simple format, which has its advantages. One advantage is that it allows measuring the lexical knowledge of low proficiency learners, as well as young learners. Another advantage of these checklist measures is that they can be constructed with relative speed and ease, as long as a suitable frequency list is available. Also a large number of words can be tested relatively quickly at least compared to other tests. (Beglar & Nation, 2014; Nation, 2013; Milton, 2009)

<i>Tick the words you know the meaning of, e.g. milk ✓</i>			
gathering	forecast	descent	revenge
strap	conscious	wodesome	heartless
untamed	mudge	topical	mere
loyalment	crope	robber	awkward

Illustration 1. Example of a Yes/No test item (Beglar & Nation, 2014:173)

The simplicity of the test format does not only have advantages, certain reservations have been voiced by researchers, such as Read (2000), that words might just be recognised by the test takers but not (fully) understood. This includes the fact that learners might be able to guess the answers to a certain degree. Despite the criticism, the Yes/No test is still in extensive use and continues to be an important format for L2 vocabulary testing. (Beglar & Nation, 2014; Milton, 2009)

5.1.2. Vocabulary Levels Test

As mentioned above, one purpose of language tests is to determine areas in which learners have difficulties or to determine whether learners know enough vocabulary for particular tasks (diagnostic test). This can happen by making use of a rather popular test method, which requires test takers to show their vocabulary knowledge by matching translations or explanations of words. These tests can be passive by providing translations or explanations (e.g. paraphrases or synonyms) for the test takers to choose from (i.e. *recognition tests*) or productive by providing a L1 stimulus for the test taker to provide a foreign language word (i.e. *recall tests*). The *Vocabulary Levels Test (VLT)* (Nation, 1983; Schmitt, Schmitt & Clapham, 2001) is an example of a recognition test because learners are provided with target words and explanations to match to these words, as can be seen in Illustration 2. This allows not only

testing for (passive) recognition of word form but also estimating knowledge of words and their meanings. The words contained in the VLT are words of four frequency levels (measuring knowledge at the 2,000, 3,000, 5,000 and 10,000 word frequency levels) as well as a section for academic words. One factor that plays a role is that this is a rather complex test, which means that the outcome of the test relies not only on the test taker's knowledge of the target words (on the left in the example below) but also on the knowledge of the words used in the explanations (on the right in the example below). (Milton, 2009; Nation, 2013)

1. original	
2. private	_____ complete
3. royal	_____ first
4. slow	_____ not public
5. sorry	
6. total	

Illustration 2. Example of a Vocabulary Levels Test item (Beglar & Nation, 2014:174)

The original purpose of this test was to see on which (level of) vocabulary learners should focus on. The vocabulary could be high- or low-frequency words or academic vocabulary. This is the reason for the total score not being of primary interest but the scores on the individual frequency levels being far more interesting. When taking the Vocabulary Levels Test, the test takers are not only required to match words with short definitions or translations, the test also contains distractors, which do not have semantic relationships with the correct response. Consequently this test is sensitive to partial knowledge of the target words. Similar to any other test, however, it is not possible to test every single word in a particular group but the test still needs to be representative, which is the reason item selection during test design has to be considered very carefully. (Beglar & Nation, 2014; Nation, 2013)

5.1.3. Vocabulary Size Test

As the name of the test already suggests, it is used to measure vocabulary size of both L1 and L2 speakers of English. The *Vocabulary Size Test (VST)* was created by Paul Nation and David Beglar (2007) with the help of sampling from word family lists, which were developed from the British National Corpus but now also include data from the Corpus of Contemporary American English (COCA). The VST measures vocabulary knowledge on different word frequency levels up to the fourteenth or the twentieth 1,000 word family level. This shows that compared to previous tests the VST extends the range of measurement. The test is designed in a multiple-

choice format but different from the Vocabulary Levels Test the correct answer and the distractors share elements of meaning, which makes the VST more difficult than the Vocabulary Levels Test. The only way in which the version testing up to the fourteenth frequency level and the version testing up to the twentieth frequency level differ is in how many test items they contain. The former contains 140 multiple-choice items, 10 for each 1,000-word-family level, while the latter contains 100 multiple-choice items. Both versions of the VST test the same kind of knowledge, which is written receptive vocabulary knowledge. This includes written word form, the connection between form and meaning and concept knowledge, at least to a small degree. Consequently the score on this test does not indicate how well the test taker can use these words when speaking or writing. The target word in the VST is largely decontextualized but still appears in a short and non-defining syntactic context in the test, which only gives information about which part of speech the target word is used as. Apart from being available as an English monolingual version, bilingual versions of the test have also been developed in quite a number of languages. This includes Chinese, Japanese, Korean, Vietnamese, Thai, Arabic and Russian. The reason for this is that the bilingual versions of the VST are more suitable for low proficiency learners. Nation (2013:524) includes an example of both the English monolingual and the Chinese bilingual version, which can be found in illustration 3. (Beglar & Nation, 2014; Nation, 2013; Nation & Beglar, 2007)

soldier: He is a soldier. a. person in a business b. student c. person who uses metal d. person in the army	soldier: He is a soldier. a. 商人 b. 学生 c. 金属工艺制造者 d. 士兵
---	--

Illustration 3. Examples of Vocabulary Size Test items (Nation, 2013:524)

The results of the VST can be used for either practical purposes of language teaching or for research purposes. Practical implications of test results would be for designing the syllabus, selecting reading materials and planning vocabulary instruction. Research purposes on the other hand can be measuring the total receptive written vocabulary size of both native speakers and language learners of a certain age or educational level. (Beglar & Nation, 2014; Nation, 2013)

5.2. Assessing vocabulary depth

As discussed in chapter four, word knowledge does not consist of only a single quality such as recognising word forms and attaching meaning to them. Language users have to know other things about words in order to use them completely in communicative contexts. Nation (2013:49) distinguishes 18 different qualities of word knowledge. This is captured by the dimension of vocabulary depth. It is important to add, though, that breadth and depth are only seemingly independent from each other. Rather, they have a relationship, which is, according to Beglar and Nation (2014:178), considerably consistent. It is a correlation, which changes according to participants' proficiency levels: the correlation is lower for low proficiency learners than it is for high proficiency learners. (Beglar & Nation, 2014; Milton, 2009)

The aspect of vocabulary depth can be difficult to measure because it is difficult to define it without ambivalence. This is also a challenge to the validity of any test of vocabulary depth. Because “[w]ithout a clear construct, it is impossible to create a test that can accurately measure and quantify a quality whatever that quality is.” (Milton, 2009:150). The idea of a dimension of vocabulary that might give us information about the degree of accuracy, appropriateness or native-likeness of language learners, which is separate of the dimension of vocabulary breadth, persists. Thus, different approaches to assess vocabulary depth have been developed by making use of different measures. One approach of a qualitative nature is to interview language learners about the information they know about certain words. According to Nation (2013:552) the advantage of interviewing is that the researcher is more flexible concerning what is tested (i.e. multiple dimensions of vocabulary knowledge at once or one dimension in greater depth). Another approach that has become increasingly popular is word association tasks, which have been used in psychology for a long time. Generally speaking, when this type of task is used (in a productive manner), the test taker is given a target or stimulus word (e.g. *white*) and then has to produce a word in response. In the case of our example, the word *white*, *black* would very likely be a response. An advantage of word association tasks is that a larger number of low frequency words can be elicited in a shorter amount of time than by using writing or speaking tasks. The reason for this is that in free production (i.e. writing or speaking) a large amount of high frequency words is necessary for well-structured, grammatical language and therefore it would take a high quantity of language production until a suitable amount of low frequency words were used to gain meaningful results. (Milton, 2009; Nation, 2013)

Although other tests for vocabulary depth exist, in this section we will discuss only the three most well known tests in this area: the self-assessment vocabulary knowledge scale, the word associates test and the so-called V_Links measure.

5.2.1. Vocabulary Knowledge Scale

Wesche and Paribakht (1996) developed the *Vocabulary Knowledge Scale (VKS)* in order to track the development of knowledge of new words. The test takers are presented with a list of target words and a rating scale, which includes five points for the test takers to rate their own knowledge of the target word. The scale's five points represent the depth to which the target word is known combining selected and constructed responses, which can be seen in Illustration 4 below:

- I. I don't remember having seen this word before.

II. I have seen this word before but I don't know what it means.

III. I have seen this word before and I think it means _____.
(synonym or translation)

IV. I know this word means _____. (synonym or translation)

V. I can use this word in a sentence: _____ (write a sentence)

Illustration 4. Vocabulary Knowledge Scale (Beglar & Nation, 2014:179)

The advantage of the VKS is that it includes both receptive and productive aspects of learners' vocabulary knowledge with the assumption being that the more deeply learners know a word the easier they are able to fill out the productive parts of the scale. But although researchers in the L2 vocabulary area have used the VKS quite frequently (e.g. Horst & Meara, 1999; Milton, 2008), many questions about its functioning and validity remain. Criticism has been voiced on the insensitivity of the test concerning many aspects of depth of vocabulary knowledge because multiplicity or shades of word meaning cannot be tested, for example. Number five on the scale, using the word in a sentence, was also criticised because this kind of productive use allows the test taker to use semantically neutral contexts for the target word to be placed in, which gives little clue about whether the test taker also really understands its meaning or not. Evidence for this can be found in a study with Hong Kong students, who were quite capable of producing correct sentences with the target word, also when they did not know its meaning at all (Mcneill, 1996). This also shows that the middle of the scale is of almost no use and the scale becomes somewhat of a binary test: I know this word / I do not know this word). As we can see in the box above, this mix of receptive and productive knowledge is quite a big concern. It is possible,

however, to reformulate the VKS into separate measures. The original purpose of this scale is to track the development of word knowledge of language learners, which this scale can definitely do by tracking changes in the ratings of the scale for the same target words over time. (Beglar & Nation, 2014; Milton, 2009)

5.2.2. Word Associates Test

Because the self-assessment scale of the VKS has proved insufficient, we have to look elsewhere for a test that measures a variety of qualities in one test so as to provide better coverage of this dimension. The *Word Associates Test (WAT)* developed by Read (1993, 1995, 2000) is the most well known test in this area. The theoretical base for the WAT is the study of word associations. The WAT should not be confused with the word association tests used in psychology. While the psychological word association tests ask the test taker to respond with the first word or associate coming into mind when hearing a certain word, the WAT is interested in which words the test taker thinks go together (i.e. form collocations) or have similar meaning (i.e. are synonyms). For better understanding we shall look at the theoretical background shortly. Generally speaking, there are two main types of associations: syntagmatic and paradigmatic associations. As stated by Meara (2009:6) syntagmatic associations complete a phrase (or syntagm), as is the case with *brush* and *teeth* (see the discussion on collocations in chapters two and three), while paradigmatic associations belong to the same part of speech and share meaning: *tree* and *bush*, for example. Additionally, there are also associations where words share phonological similarities, which are more common among children and beginner level second language learners. This type of association is called *clang associates* and examples are *blue* and *glue*, as well as *night* and *knight*.

The initial version of the WAT differs from the version we find today because it was done via an interview. Read combined a selection of words and open-ended questions, which were designed to test aspects of the learner's vocabulary knowledge. A typical test item of this version is shown in Illustration 5 below, which attempts to assess the learner's knowledge of different shades of word meaning for the word *interpret* (including collocations, associations and derivations). But even this detailed approach to assessing depth of vocabulary knowledge does not manage to include or address every aspect of it. Furthermore, this approach is rather time consuming while being able to test only a small number of words. (Milton, 2009; Read 2000)

TO INTERPRET	
1. Write two sentences A and B. In each sentence, use the two words given. A interpret experiment B interpret language	
2. Write three words that can fit in the blank. to interpret a(n) _____. i _____ ii _____ iii _____	
3. Write the correct ending for the word in each of the following sentences. Someone who interprets is an interpret_____. Something that can be interpreted is interpret_____. Someone who interprets gives an interpret_____.	

Illustration 5. Example of a Word Associates Test item (Milton, 2009:162)

Because the initial version seemed rather complicated, Read (1993) tried to simplify the testing procedure consequently making it quicker. The revised version of the WAT requires learners to select responses from a given sample. It presents a target word with eight options, four of which are semantically related to the target word. Illustration 6 is an example of a test item. As the words in the example show, test takers' are required to identify words that have a certain relationship with the target word *sudden*. The test takers should identify the words on the left, which have a paradigmatic relationship with the target word: *quick* is a synonym of *sudden*, for example. The words on the right on the other hand have a syntagmatic relationship with the target word, such as collocations (e.g. *sudden* and *change*). The example also shows that testing the knowledge of word parts and derivations were dropped in the development of the revised version of the test. (Beglar & Nation, 2014; Milton, 2009; Nation, 2013)

sudden							
beautiful	quick	surprising	thirsty	change	doctor	noise	school

Illustration 6. Example of a Word Associates Test item (Beglar & Nation, 2014:178)

According to research (such as Qian & Schedl, 2004) the WAT has several advantages. First of all, the inclusion of synonymy, polysemy and collocation into the WAT can be used to describe proficiency levels in vocabulary knowledge. Secondly, the fact that this test focuses on receptive vocabulary knowledge only makes it an appropriate test of the knowledge needed for reading comprehension. But of course this test format as any other has certain drawbacks as well. This is the case especially when the WAT is used for large-scale standardised vocabulary assessment. The main concern here is the development of uncontroversial answer keys, especially for collocations. This is also true for the current study, as it was not easy to find an uncontroversial answer key, so that three different answer keys were prepared. More on this will follow in chapter six on research methodology. Furthermore, the extent of vocabulary covered

by the WAT is limited because only adjectives are used as the stimulus or target words but in the collocation section of the test nouns were also tested to a certain extent. As the WAT is the test type chosen for the present study, more detail on the design and scoring of a WAT test item will be discussed in the section on methodology.

5.2.3. V_Links

Hoping to encapsulate all aspects of the dimension of vocabulary depth, this measure – called V_Links – was developed by Wolter and Meara (2004). This test tries to assess and count the number of links of all kinds between words. By looking at which associations language learners and users make, patterns of links might emerge. To illustrate this measure an example can be found below. The numbered circles are the target words used in the test, and the arrows between them are the links that this learner, a non-native speaker of English, identified. (Milton, 2009)

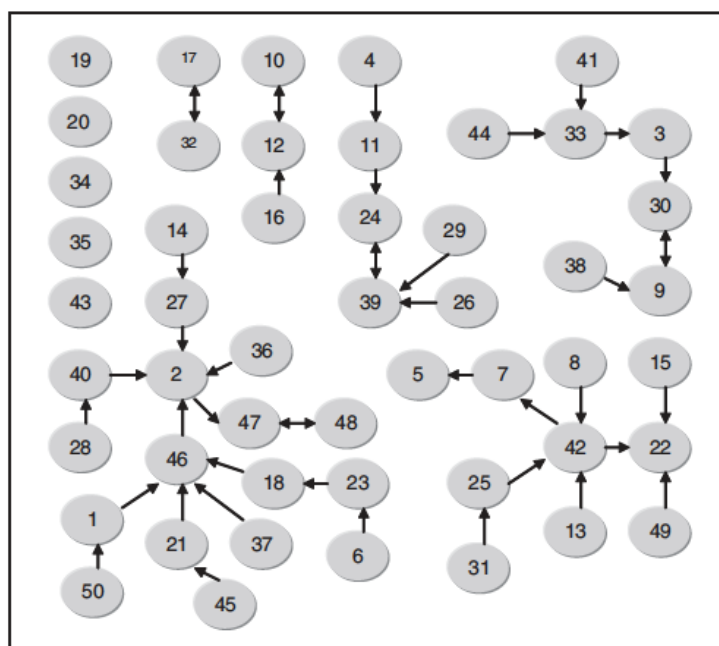


Illustration 7. Example of a V_links measure (Schur, 2007:190)

The example shows that some words (e.g. words 42 and 46) seem well connected, while others (e.g. words 34 and 35) are not. In English some words, such as the verb *get*, are highly connected. Looking at the verb *get* we can see that it not only links frequently with pronouns (e.g. *you get*, *she gets*, etc.) but also with prepositions to make phrasal verbs (e.g. *get up*, *get off*, etc.) and with noun phrases, as well (e.g. *get married*, *get divorced*, etc.). Other words that appear less frequently are more restricted in their use and collocate with a smaller number of words (e.g. *dint*, *kith*).

Of course other versions of the V_Links measure exist as well, such as Wolter's (2005). For this test 10 words are selected randomly from a wordlist, which are then presented to the test takers in a circle. They in turn then mark the links between the given words. In the example below of Wolter's (2005) test, it can be seen that a mix of different word classes (e.g. adjectives, verbs or nouns) is included. The arrows in the example show that *news* might associate with *story*, while *true* also associates with *story* but also with *feeling*. With this type of test, it has been suggested that native speakers are able to find more connections in such a set of words than non-native speakers, simply because they (probably) know more potential links between words and also because their knowledge of depth of vocabulary is much greater. But as non-native speakers enhance their language proficiency, their ability to find connections between words increases. (Milton, 2009)

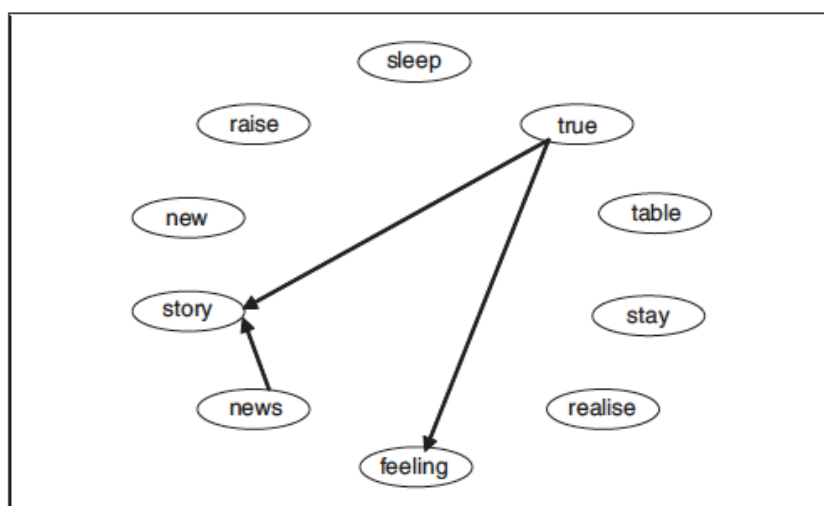


Illustration 8. Example of a V_links measure (Wolter, 2005:144)

The biggest problem with a test of this type, as described above, is that the quality of associations involved cannot be checked in any way. Which is why as soon as the learners realise the purpose of the test is to connect as many words as possible, they might start connecting words whether or not a connection or association really exists between them. Researchers are trying to solve these problems – foremost by using words of high frequency and by referring to native speakers' knowledge to compare non-native speakers to. (Milton, 2009)

This chapter has discussed several different tests for assessing different aspects of vocabulary knowledge. However, this has only scratched the surface, as many more tests (e.g. for productive vocabulary knowledge) exist. As mentioned in the introduction, for the current study

one type of test was chosen: the Word Associates Test. Full argumentation concerning test selection can be found in the next chapter on research methodology.

6. Research methods

In previous chapters we have seen the features of the lexicon in general, as well as specifically for the Chinese language. The way we ‘save’ words in our brain makes collocations interesting because there are many links between the words in our (mental) lexicon. In authentic language sequences words generally do not stand on their own but appear in combination with other words and may acquire a separate status as a “combination” through frequent co-occurrence. This makes the examination of collocations very interesting also for foreign language learning as conclusions for language teaching can be drawn from such studies. Another factor in choosing this topic has been that although there has been a lot of research on lexical second language acquisition, there are huge differences in which languages are examined. This is also true for the sub-areas of vocabulary breadth and depth. It should come as no surprise that a huge number of studies focuses on English as a foreign language (EFL) (e.g. Laufer & Nation, 1995; Meara, 2004; Nation, 2013; Webb, 2016). Comparatively little research has been done on CFL generally and lexical acquisition of learners specifically. This is the reason why this master thesis examines the depth of vocabulary knowledge of CFL learners. Due to the scope of this master’s thesis, this can only be seen as a start. This study, then, focuses on the ability of the MA students of the Chinese department at the University of Vienna to identify collocations in the Chinese language. It aims to answer the research questions found below.

RQ1. How well can the CFL learners in this study identify Chinese collocations?

RQ2. Does word frequency impact the CFL learners’ ability to identify collocations?

RQ3. Do other factors (e.g. language learning background or a year abroad in China) impact the ability to identify collocations?

RQ4. Is there a difference concerning the ability to identify collocations between the two sub-groups among the participants (i.e. first and third semester MA students; heritage and non-heritage learners)?

In terms of instrumentation, research questions one and two will be addressed via the Word Associates Test (WAT), RQ3 with the help of a demographic questionnaire and RQ4 by correlating the results from both instruments.

This chapter will discuss the following features of this study: the model study it is based on, the adapted test, and the context of this study (i.e. participants, data collection), as well as the data analysis methods.

6.1. Model study

This study is based on a model study by Chen (2016), who explored the depth of vocabulary of high proficiency learners of CFL in the USA focussing on how well the participants can identify two types of word association: synonym and collocational association in relation to word frequency. This section will describe the important aspects of this study (methods, participants, etc.) and will be put in relation to this thesis' study in the following section. Considering the large number of tests that exist for assessing vocabulary knowledge, why was this specific study chosen to be replicated? This study was chosen as the model study for a number of reasons, two of which are practicability and face validity. Additionally, the number of predesigned tests for CFL is rather small – compared to other languages, as most existing tests by far were developed for English as a foreign language.

The model study was comprised of three parts: a *Word Associates Test (WAT)*, a *demographic background questionnaire* and an *interview*. This shows that the study made use of two quantitative methods and a qualitative method. We will now look at each part of this study in the above order, starting with the *WAT*. Because the aim is to examine learners' identification of synonyms and collocations, the *WAT* is the central element of the study. Each part of the *WAT* (one part for synonyms and one for collocations) consisted of 45 stimulus or target words, which are all part of the same word class, adjectives. However, Chen did not provide an explanation as to why she included only one word class in the *WAT* and why she chose adjectives. Target words for the creation of test items were selected from three sources. One source was the *Commonly used Chinese words frequency list* (汉语常用词词频表 [Hànyǔ chángyòng cí cípín biǎo]), which was developed by the Beijing Language and Culture University based on a corpus of contemporary Chinese that contains about 500 Million Chinese characters. Word frequencies were calculated based on the occurrence of the words inside the corpus. Another source was the *Chinese synonym dictionary* (同义词大词典[Tóngyìcí dàcídiǎn]) that was published by the Shanghai Lexicographic Publishing House (上海辞书出版社 [Shànghǎi císhū chūbǎnshè]) and

contains almost 15,000 synonyms and near synonyms. The third source of words was the CCL Corpus developed by the Centre for Chinese Linguistics at Peking University including a total of 783,463,175 characters.⁹ The words for the test items were of course not chosen at random, but Chen observed the following criteria, which are listed below (Chen, 2016:55):

1. All words had to be adjectives or had to be predominantly used as such.
2. All words had to be two-character compound words (for consistency).
3. All words came from three levels of word frequency: below 1,000, 1,000 to 5,000 and above 5,000.

Choosing two-character words for both the target words and the collocations and synonyms is very reasonable and comprehensible because the majority of Chinese words fall into this category, which we saw in chapter three on the lexicon in the Chinese language. Additional guidelines were set for creating the options of the test items, some of which are listed below (for a complete list see Chen (2016:56-57)).

1. Four options of synonym and four options of collocational associations were designed for each item.
2. To minimize the impact of word frequency on the participants' responses, all options are of equivalent or higher word frequency than the target words (if possible).
3. Collocations were selected from those appearing frequently in the CCL Corpus.
4. The chosen distractors for collocations with the target word did not appear in this way in the CCL Corpus.
5. Within an item, the number of associates and distractors has some variations within the test booklet (especially between adjacent items).

In the previous chapter on the assessment of vocabulary knowledge, we discussed the WAT developed by Read (1993) and an example for a typical test item was discussed as well. While the basis for this test is indeed Read's WAT, the format of the test items was adapted by Chen. The test items for synonym associations and the ones for collocational associations were divided into two separate booklets. This differs from the typical test items we saw in the previous chapter, where all eight responses were placed below the target word in one and the

⁹ According to Chen (2016), the data for the corpus was collected from oral and written linguistic materials in modern day use (e.g. modern literary works, applied writing, transcripts of TV shows, newspaper articles, etc.).

same item. An example of a typical test item in the Chinese WAT can be found in the next section on the adapted test.

The development of the WAT further contained two pilot studies: one with Chinese native speakers and one with CFL learners. First, the pilot study with Chinese native speakers was conducted. Initially, the WAT contained 70 items and each of these items tested the two types of word association simultaneously, which shows that initially the test items were modelled directly on Read's item type. In this phase of the pilot study a true/false format was used for the Chinese native speakers to decide which combinations were word associations and which were not. In order to improve the test, feedback of the native speakers on the test items was also gathered (e.g. on the clarity of directions, the item format, difficulty and ambiguity). Words that the native speakers were less familiar with were removed from the WAT because it was likely that these words would be even more difficult for the CFL learners. Following this initial pilot study, the 70 evaluated items were split into synonyms and collocations and two booklets with 40 items each were created. After this revision of the WAT, a pilot study among CFL learners was conducted and the WAT was supplemented by a 10-minute interview on the test-taking experience. Suggestions made by the participants of the pilot study were considered in the revision of the WAT before the conduction of the main study. After revising the test once more and adding a few more test items, a review of the test booklets by a Chinese professor was also included in the pilot study phase. Target words found problematic by the professor were removed from the booklet.

Additional information on the participants was gathered with the help of a demographic questionnaire. This questionnaire collected information on the participants' language learning background, Chinese learning experience and proficiency test scores. A combination of multiple-choice and short-answer questions was used for the format of this questionnaire. The study was further complemented by interviews. The consent for taking part in an interview was asked in the last question of the questionnaire. The interviews were guided by a list of questions, which focused on the participants' impression of the test items and test administration. In addition, the interviews also sought to find out which processes and strategies the participants used to complete the WAT.

The participants of Chen's study were found by means of convenience sampling, making use of the following two criteria to screen potential test-takers. Firstly, participants had to be enrolled in at least one advanced-level Chinese language course at the time of participating, having previously learned Chinese for at least four years. The second criterion was that participants

had to be non-heritage learners of Chinese, which excludes anyone who grew up speaking Chinese at home with their relatives. By applying these two criteria, Chen initially recruited 25 participants for her study. Unfortunately, in a later screening process of the filled-out tests, eight participants had to be excluded. The reason for this was that it turned out that they were heritage-learners after all or had not learned Chinese for a long enough period of time or did not indicate their length of study at all. So in the end 17 CFL learners made up the actual sample of the test, six of which took part in the interview following the test.

The collection of the data was impacted by the fact that the participants were in different locations. Consequently data collection was either done on site or through mail. As mentioned above, the WAT consisted of two booklets and the participants were required to finish both booklets within 50 minutes in consecutive order (booklet 1 had to be handed back to the proctor before starting booklet 2). After finishing both booklets, the participants filled out the demographic questionnaire. The six interviews were done individually via Skype, each one lasting approximately 10-15 minutes.

We now turn to the data analysis procedures used in Chen's study. A key part of this obviously is the scoring of the WAT, which was done in the following way: if an associate or distractor of an item was correctly identified, participants received one point. So the score for any item in the test ranged from zero to four. Incorrect selections were not penalised, while unknown words were recorded but not considered for scoring. For the comparison of the performance of CFL learners' across association types and by word-frequency levels, descriptive statistics were used: total score, rate of correctly-identified associates and correctly-identified distractors; this included mean, standard deviation, minimum and maximum.

Examining the outcomes of Chen's study, the results show that the participants did not know an average of 16 and 16.29 target words for booklet 1 and booklet 2 respectively. However, these numbers should be the same, as the target words were the same for both booklets. Therefore, a total average of 63.6% (approximately 28 target words) should be known by the participants. In reality these numbers varied widely (from a minimum of 17 to a maximum of 41). Regarding the correct identification of the synonyms and collocations, an average of 31% (55.12 points) of synonyms and an average of 37% (65.94 points) of collocations were identified correctly. However, again large variations can be found in the scores of different participants: ranging from a minimum of 24 and maximum of 105 on booklet 1 and a minimum of 35 and a maximum of 134 on booklet 2. The rate of correctly identified distractors was 30% in booklet 1 and 31% in booklet 2. Showing that in booklet 1, participants were equally competent to identify synonyms

and distractors, while in booklet 2 the participants had an easier time identifying collocations and not distractors. The additional interviews with six of the participants (Chen, 2016:83-84) also showed that the participants had different opinions on which of the booklets was harder and also the reasons for choosing which booklet was harder. For example one participant found booklet 2 easier than booklet 1 because he was more familiar with the word options included in it. Another participant found booklet 2 easier because he was able to recognise more words due to the context provided by the collocations. One participant said that booklet 1 was easier because the word options shared characters with the target words.

Chen's results also show that the number of unknown words increased with decreasing word frequency: an average of 1.5 unknown words at the high-frequency level, six unknown words at the mid-frequency level and nine unknown words at the low-frequency level (i.e. 11%, 40% and 60% respectively). The average number of correctly identified synonyms and collocations also dropped when the target word frequency dropped (for synonyms: from 51% at the high-frequency level to 27% at the mid-frequency level and 20% at the low frequency level; for collocations: from 59% to 39'8% to 23% at the different frequency levels).

6.2. Adapted test

This thesis looks at the depth of vocabulary knowledge of the MA-level students at the Chinese Department of the University of Vienna. This section will discuss how the above model study was adapted for this purpose, after obtaining permission from the author via LinkedIn-Messages to use her test in a partial replication study.

Similar to Chen's (2016) study, this study also uses a combination of measures. These include a WAT and a demographic. After consultation with the supervisor of this thesis, who also happens to be the study programme leader of the Chinese department, Professor Dr. Weigelin-Schwiedrzik, the WAT of the model study was adapted in order to better suit the (potential) participants of the study. Instead of testing for both synonym and collocational associations, this WAT only includes collocational associations. The reason for this is that the language teaching at the Chinese Department does not put a lot of focus on the learning of synonyms, which is why the students do not develop a lot of knowledge in this area and would have difficulties with this part of the test. Furthermore, some of the stimulus words in the synonym section are important for understanding literary texts but the Vienna MA programme is focused on understanding academic texts. Collocations on the other hand are part of the Vienna language-teaching programme, which is quite important to the department because the students should

experience authentic language use. The full version of the WAT of this study can be found in Appendix A but for more clarity we will look at an example for a typical test item taken from Chen (2016:147) here.

干净（的）			U
意义	Y	N	U
食物	Y	N	U
成果	Y	N	U
房间	Y	N	U

Illustration 10. Typical WAT item (Chen, 2016:136)

This test implements a multiple-choice format with three response options: Y (Yes), N (No) and U (Unknown). If the participants do not know a certain word combination (i.e. the collocation), they can mark the box U (Unknown). This is also true if the headword itself is unknown: If the participants do not know a certain headword, they can also mark the box U (Unknown) next to it and skip the entire test item. After clarifying the answer mechanism, let us consider the words in the above example in order to show how the participants should mark the test items: the target word is 干净 [gānjìng] (= ‘clean’). Four options to choose from are available: 意义 [yìyì] (= ‘meaning’), 食物 [shíwù] (= ‘food’), 成果 [chéngguǒ] (= ‘achievement’) and 房间 [fángjiān] (= ‘room’). Choices two and four (i.e. ‘food’ and ‘room’) would be marked Y (Yes) because they form expressions with the target word, such as ‘clean food’ and ‘clean room’. Choices one and three (i.e. ‘meaning’ and ‘achievement’) on the other hand would be marked as N (No), as they have no collocational association with the target word.

In total the WAT contains 45 target words, as did booklet 2 of the model study. Although Chen (2016:75) found that item 12 (the target word 简单 [jiǎndān] = ‘simple’, ‘uncomplicated’) had identical word options for both booklets: both booklets contained the same synonyms and distractors. For this reason Chen excluded this item from her data analysis and only 44 items remained.¹⁰

¹⁰ The booklet for this thesis’ study still contains item 12 in order to have an additional distractor included in the test. However, this item will also be excluded from data analysis.

This study also includes a demographic questionnaire to gain background information in order to be able to better classify the results of the WAT. The format also includes multiple-choice and short answer items but compared to Chen, this questionnaire was expanded and contains a few more questions. Besides language background, Chinese learning experience and proficiency test scores, this questionnaire also includes age, gender and what other languages apart from Chinese the participants have learned (e.g. in school). The full demographic questionnaire can be viewed in Appendix B. Interviews were excluded from this study due to the fact that interviews, their transcriptions and analysis are very time consuming and the scope of this thesis is limited. Interviews could only be included in a larger thesis. Also the interviews mostly gave feedback on the test but did not provide more information on the participant's vocabulary depth.

In summary it can be said that this study is replicating Chen's (2016) study, only partly however in order to better fit the context of the current study. Thus, only part of Chen's WAT was integrated into this study (i.e. the test booklet on collocations) and the demographic questionnaire was expanded with three additional questions.

Because the Department of Chinese studies expects its students to reach certain proficiency levels, as will be discussed in the next section, it was deemed necessary to check into which HSK and CEFR proficiency levels the words commonly fall into. Thus, the words were looked up in the HSK materials, which are sorted according to proficiency level. All target words were checked against HSK vocabulary lists to see which proficiency level they are allotted to. First of all it should be noted that it seems to be the case that Chen (2016) did not consider the proficiency level factor when compiling the test, as no system concerning the HSK proficiency levels can be detected after consulting HSK vocabulary list materials. The evidence for this can be drawn from the fact that 16 out of 45 target words are not contained in any of the vocabulary lists consulted. The majority of the other words belong to the vocabulary lists of HSK levels 5 and 6 (i.e. 22 out of 29 words), while two words belong to HSK levels 4 and 1, respectively, three words belong to HSK level 3. No words in the WAT belong to HSK level 2. An overview of this distribution can be found in Table 4 below.

HSK level	Number of target words
1	2
2	0
3	3

4	2
5	11
6	11
No level	16

Table 4. WAT target words and HSK levels

Additionally, it should be also noted that sixteen of the 45 target words are not included in any of the HSK vocabulary lists. These can be found in all three frequency levels, however, the low-frequency word section contains the most words that are not listed in the HSK vocabulary material (i.e. nine out of 15 words of this section are not attributed to a HSK level). The high-frequency section contains one word and the mid-frequency section contains six words that are not contained in the HSK vocabulary lists. This suggests that Chen (2016) did not consider language proficiency (in form of the HSK vocabulary material) when designing this WAT.

6.3. Context of the study

This study focuses on the depth of vocabulary knowledge of students at the Chinese Department of the University of Vienna. The department offers a three-year BA programme in Chinese Studies and one MA programme with two different foci. The BA programme includes the following foci: language, history and society, literature and culture, as well as politics, law and economy. The MA programme offers the following two branches: the “science branch”, which continues focusing on the foci mentioned above, and the “teaching branch”, which focuses on linguistics, didactics and pedagogy preparing graduates to teach Chinese as a foreign language.

After consulting the supervisor on which students would be suitable for the WAT, it was decided that the best option was to recruit students doing their Master studies due to the fact that these students had been learning Chinese longer and therefore have a higher proficiency level. In total 16 participants were recruited for this study; six of them were first semester MA students and twelve were third semester MA students. This opens up the possibility for a comparison between these two groups of students to evaluate if there is a difference between the scores of these two groups and whether a development of this specific language ability can be seen in students from the start of the MA programme and the end of the programme.

In order to better contextualise the vocabulary focus of the present study the language component of the BA and MA programmes will be described in more detail. Practically all students who are not heritage language speakers of Chinese join the department with zero

competence in the Chinese language. On the BA level, therefore 90 of the programme's total 180 credit points (i.e. 50%) are dedicated to language acquisition. Language courses are spread across all six semesters, with four courses each semester: from the basics in the first semester to individual grammar, writing, speaking and reading courses to reading press, specialist language courses (e.g. Business Chinese) and even Classical Chinese in later semesters. Table 5 below shows the language courses and credit distribution in detail.

Semester	Course title	Credits
1	Modern Chinese 1a	5
	Speaking Lab 1a	3
	Speaking Practice 1a	3
	Theory and Practice of Chinese Writing and Language	4
2	Modern Chinese 1b	5
	Speaking Lab 1b	3
	Speaking Practice 1b	3
	Written Chinese 1b	4
3	Modern Chinese 2a	5
	Reading Practice 2a	4
	Speaking Practice 2a	2
	Written Chinese 2a	4
4	Modern Chinese 2b	5
	Reading Practice 2b	4
	Speaking Practice 2b	2
	Written Chinese 2b	4
5	Listening and Speaking A	2
	Writing and Reading Chinese A	5
5/6	Classical Chinese	6
	Specialist Language	5
	Reading Press	5
6	Listening and Speaking B	2
	Writing and Reading Chinese B	5

Table 5. Language courses and credit distribution of the BA programme

Teaching materials also vary from semester to semester with textbooks being used in every course until the fourth semester. From then on some courses (such as Classical Chinese or Business Chinese) use textbooks, while others (such as reading and writing) do not but use teacher-designed materials provided via an e-learning platform.¹¹ Language learning in the BA programme of Chinese studies at the University of Vienna is very intensive and is designed above all to expand the students' vocabulary. The aims of the BA programme for the language proficiency of its graduates is a B2 (i.e. upper intermediate) level according to the Common European Framework of References for Languages (CEFR), which would translate to level 4 according to the 汉语水平考试 [Hànyǔ Shuǐpíng Kǎoshì] (HSK), a standardised proficiency test for the Chinese language.

In the main MA programme language learning is organised differently. In total there are four language courses, which amounts to one course per semester. These courses concern both written and oral translation practice, with two courses each. For oral translation there is no textbook but audio recordings. In written translation the same textbook is used for both courses.¹² The aim of the master programme for the language proficiency of its graduates is C1 (i.e. effective operational proficiency or advanced) according to the CEFR, which would translate to level 6 according to the HSK.

6.4. Data collection and analysis

The collection of the data was done in a printed, paper and pencil version during the sessions of the participants' language courses. The test booklets were administered by the researcher who made sure no helping material, such as dictionaries, were used. The WAT and the demographic questionnaire were printed in one booklet and handed out at the same time. For the WAT a time allotment of 25 minutes was given, which is half the time of the model study but it was only half

¹¹ In the first two semesters the textbooks are: Xia & Zeidl (in preparation) 在汉学系学汉语 (上 / 下) [zài Hànxuéxì xué Hànyǔ (shàng/xià)] *Learning Chinese at the Chinese department (Parts 1 and 2)*. In the third and fourth semester the textbooks are: Yang (ed.) (2006) 汉语教程 第二册 (上 / 下) *Hànyǔ Jiàochéng dì èr cè (shàng/xià) Chinese course 2 (Parts 1 and 2)*. Beijing: Beijing Language and Culture University Press.

¹² The textbook is: Li & Weigelin-Schwiedrzik (2007) 桥—多功能汉语读本 *Qiáo – duō gōngnéng hànyǔ dúběn. Bridge – Multifunctional Chinese Reader*. Shanghai: Shanghai University Press.

of the original WAT. After these 25 minutes the participants had additional time to fill out the demographic questionnaire but most finished both within the time allotted to the WAT.

In order to better categorise the results, the test booklets were then coded with P1 through P6 for group one (i.e. first semester MA students) and P7 through P16 for group two (i.e. third semester MA students).

Naturally, any collected data has to be analysed in order to gain results. Concerning this study the first step in the analysis procedures is the *scoring of the WAT*. This was done in the same manner Chen (2016) used in her study: one point was awarded for correctly identified associates or distractors of an item. This means that the score of any given item can range from zero to four. Incorrect choices are not penalised and the unknowns are not considered for scoring but are recorded.

It should be noted here that in the beginning of the scoring process the original solutions to the WAT were not accessible to the author of this thesis. Consequently, the author found her own approach to scoring the participants' WAT results. Three different solution sheets were worked out with the help of the CCL corpus. The first solution regards every word option appearing together with the target word (no matter how frequently) as a collocation, which is a very broad interpretation of the notion collocation. The second solution excludes all word options that appear together less than ten times in the CCL corpus (i.e. all word options appearing together with the target word more often than ten times are regarded as collocations), which is a more narrow interpretation of collocation. The third solution is an even more narrow interpretation of collocations, as every word option appearing less than 50 times in the corpus is excluded from being a collocation. Thus, the results will also be compared regarding the three solutions. Depending on which scoring solution yields the highest results, an inference can be made about the type of collocation the participants have learned (more broad or more narrow collocations). After scoring the results of the WAT by making use of the three different solution sheets, the original solution was made available by Chen and thus was included in the scoring process in order to be able to better compare the results of the current study with the study it is modelled on.

In the next step of analysis *descriptive statistics* was used to answer the first two research questions (i.e. identification of collocations and differences concerning word frequency). In order to do so three indices were used: total score, rate of correctly identified associates and rate of correctly identified distractors. Correctly identified associates and distractors were evaluated

separately because it is interesting to see whether the participants can identify collocations or distractors more easily. The descriptive statistics for the indices above include average, median, minimum and maximum. To answer RQ3 the results of the WAT are looked at in connection in order to identify possible influencing factors. RQ4 is answered by comparing the results of the two groups of participants.

7. Results

The aim of this study was to investigate how well CFL learners knew Chinese collocations, whether that knowledge varied by word frequency level and which other factors might have an impact on that knowledge. The test data was analysed regarding number of known and unknown target words or word options, as well as regarding the number of accurately identified collocations.

This chapter contains the results of the data analysis in two sections. The first section presents the results of the demographic background questionnaire. The second section on the results of the WAT is divided into three parts: first the identification of collocations, secondly their relation to word frequency and third differences between the two groups of participants (i.e. first and third semester students; heritage and non-heritage learners). The discussion of the results and conclusions drawn from them will follow in chapter eight.

7.1. Demographic questionnaire

This section will describe the results of the demographic questionnaire of the whole group of participants, which was conducted in order to be able to better understand the results of the WAT. The results of the demographic questionnaire will be shown question by question, starting with the questions regarding the participants' age and gender (Q1 and Q2 in the questionnaire). The mean age of the participants in this study is 29,6 years with a minimum of 20 years and a maximum of 71 years. Eleven of the sixteen participants are 24 to 26 years old (68.75%). Second, the majority of the participants (69%) were female, while only five participants were male (31%).

The L1 of the participants is our next focus (Q3): the majority of participants listed only one L1 (14 participants or 87,5%), while two participants listed two languages as their L1. These two participants listed their L1 as German and Cantonese and German and Chinese, respectively.

Excluding these two participants with more than one L1, the languages the participants stated as their L1 include German (50%), Chinese (21.4%) and Slavic languages (i.e. Russian, Polish and Slovenian) (21.4%) and English (7.1%). The next question is also related to the question of L1 (Q4): asking the participants whether they had any family members who speak Chinese as their native language. Seven out of sixteen participants (43.7%) answered this with 'yes', five of which listed Chinese as their L1 or one of their L1s. An additional question asked those participants who answered Q4 with 'yes' how often they speak Chinese with their family members (Q5). 57.1% of them answered 'almost always', 28.6% 'sometimes' and 14.3% 'most of the time'.

Because the participants are all learners of CFL, naturally a question of how long they had learned CFL so far was included into the questionnaire (Q6). Figure 1 shows that five of the participants have learned CFL formally in a classroom for four to five years, while three participants have learned CFL for three to four years, two participants have learned CFL for more than five years, while two participants have learned CFL for less than two years and for two to three years, respectively (with also two participants not answering the question).

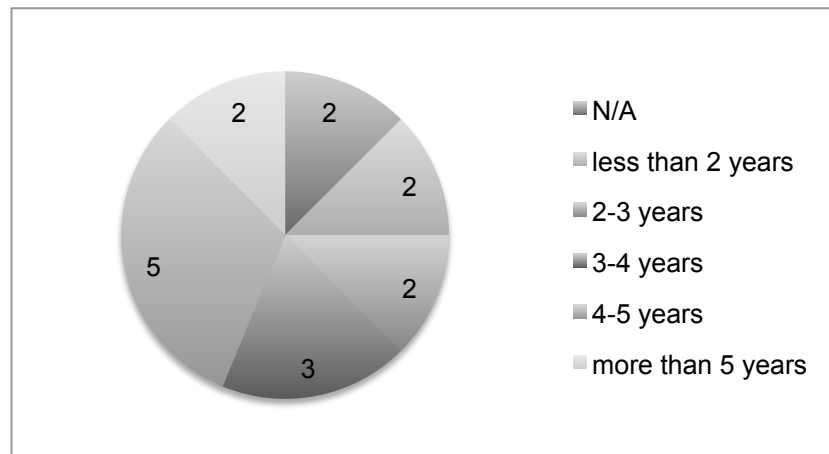


Figure 1. Time spent learning CFL

Relating to this, Q8 asked the participants whether they also learned Chinese outside the classroom from native speakers. The majority of participants (68.75%) answered this question with 'no', while 6.25% did not answer this question. 25% of participants answered that they did learn from native speakers outside the classroom: 50% of these respondents attended Chinese schools at the weekends where they learned Chinese from native speakers, 25% learned Chinese from their neighbour and 25% did not specify in what context they learned Chinese outside the classroom.

Another question designed to better understand the participants' Chinese language learning trajectory is the following, asking whether they completed their BA degree at the University of Vienna or not (Q7). 50% of the participants completed their BA degree in Chinese Studies at the University of Vienna, 31.25% answered this with no (of which 40% answered that they have not yet completed their BA degree at the University of Vienna). In case of the answer to this question being 'no' the participants were asked to share where they completed their BA degree. This includes the University of Tübingen, as well as the Free University of Berlin in Germany and another degree completely (i.e. Slavic studies). However the person answering they studied Slavic studies before did not explain where. Additionally, 18.75% did not answer this question at all.

As discussed in chapter four (section 4.4. *Important factors for learning vocabulary*), staying in a country where the language is the native language naturally aids the learning of this language. Thus, a question was included in the questionnaire whether the participants have lived, studied or worked abroad where Chinese is spoken as a native language and if they did how long they stayed there (Q9). 13% of the participants stated they have never lived, studied or worked abroad where Chinese is a native language. The majority (87%), however, said they did. The duration of these stays varied. As shown by Figure 2 44% stayed three to six months and 25% stayed seven to twelve months, 6% stayed two years, six years or answered with "not countable".

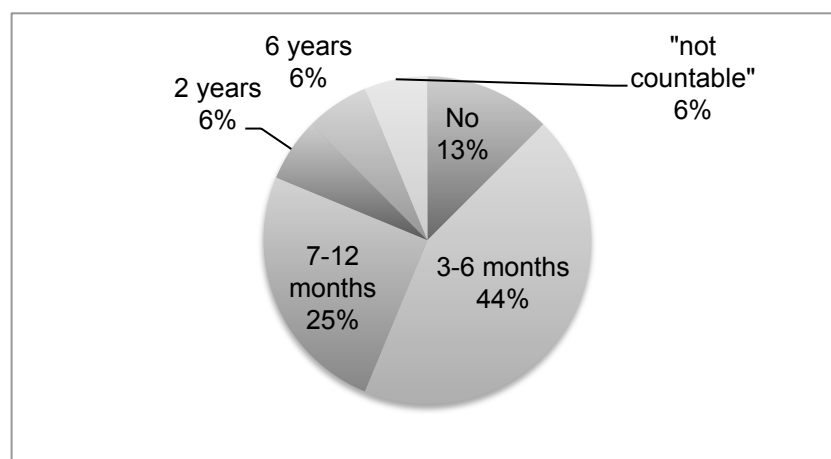


Figure 2. Time spent in a country with Chinese as its native language

Language proficiency tests are also interesting in the context of learning a foreign language. Consequently, such a question was also included in this questionnaire (Q10). Only 31.25% of the participants said they have taken a language proficiency test of Chinese (e.g. HSK or TOCFL) before, while 62.5% not answered the question, which may be because the options for

answering did not include the option 'no'. However, 6.25% still wrote down 'no'. All of the participants who answered with 'yes' have the HSK (汉语水平考试 [Hànyǔ shuǐpíng kǎoshì]) (levels four to six).

The last question of the questionnaire (Q11) asked the participants which other languages (apart from Chinese) they have learned. The questionnaire listed six options (i.e. English, French, Spanish, Italian, Japanese and Korean) and the option 'other' (i.e. to write down other languages they have learned) for the participants to choose from (Note: multiple choices were possible). 43.75% of the participants have learned two foreign languages, 18.75% learned three foreign languages, 12,5% have learned one and four languages respectively, while 6.25% have learned more than four foreign languages (apart from Chinese). Figure YZ below shows the foreign languages that the participants stated to have learned (apart from Chinese). A total of thirteen different languages can be counted.

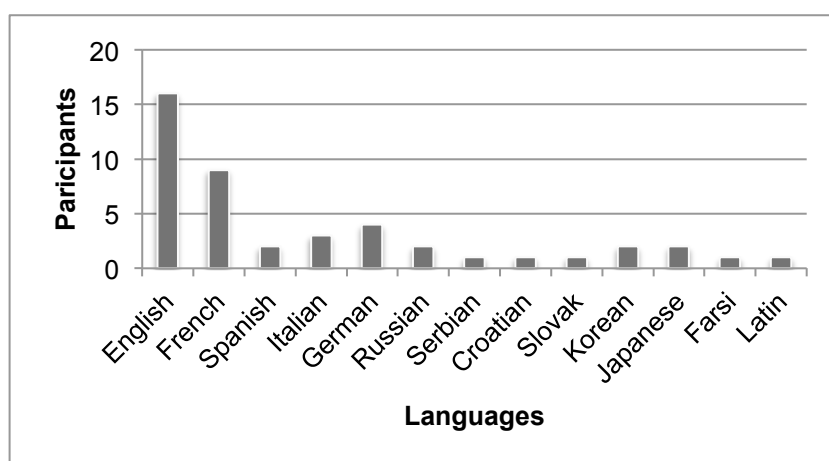


Figure 3. Languages learned by the participants

As can be seen in Figure 3 above, all of the sixteen participants answered that they have also learned English. The second most learned language is French followed by German and Italian. Only two participants stated that the only foreign language they have learned (apart from Chinese) is English.

7.2. Word Associates Test (WAT)

This section will discuss the results of the Word Associates Test (WAT) conducted for the present study. The results will be discussed in three steps: the identification of collocations and distractors in general, the identification of collocations in relation to word frequency and whether

there are differences to be found between different participant groups (i.e. first and third semester students, heritage and non-heritage learners).

As discussed in chapter six on methodology, a total of four different solutions were worked out (three by the author of the current study and one by the author of the model study, Chen (2016)). However, the results discussed here will only compare three of these solutions, as one was rendered insignificant. The solution key that was excluded was the one, which counted as collocations only those combinations, which appear in the CCL corpus more than ten times. The reason for this is that this solution key and solution sheet one (the solution, which considered every word combination appearing together in the corpus as collocations) produced rather similar results. Thus suggesting that this difference in frequency level does not impact the results of the WAT tremendously. Solution sheet three considering every word combination appearing more than fifty times in the corpus is included in this discussion because the results were very different from solution sheet one (i.e. every word combination appearing together in the corpus is a collocation) and solution sheet four worked out by Chen (2016). Accordingly, solution sheet three will be called solution sheet two from this point forward, as only three solution sheets are used and one was excluded. For better understanding of the results in chapter seven, the solution sheets will be summarised here and in Table 6 below, as the results of three solution sheets will be compared. Solution sheet one regards all combinations appearing together in the CCL as collocations, solution sheet two regards all words appearing together more than 50 times in the CCL corpus as collocations, while solution sheet three makes use of the solutions provided by Chen (2016).

Solution sheet	Definition of <i>collocation</i>
Solution sheet 1	all word combinations appearing together in the CCL corpus
Solution sheet 2	all word combinations appearing together more than 50 times in the CCL corpus
Solution sheet 3	all word combinations appearing together in a collocation dictionary (solution was provided by Chen, 2016)

Table 6. Definition of solution sheets

The analysis is based on a total of 3.520 data points collected via the distribution of 16 questionnaires.

7.2.1. General identification of collocations

This section will discuss the general findings of the WAT. Before looking at the results of the participants, the possible scores of the test should be mentioned: the total score that could be reached is 176 points, with 111 collocations and 65 distractors. First, the results gained by applying solution sheet one (every word combination appearing together in the CCL corpus is a collocation) are discussed. The minimum score that was reached is 36 points and the maximum score is 147 points. The average points reached by the participants are 78.5. The median for the total score is 58.5 points. The minimum of correctly identified collocations and distractors is 23 and 10, while the maximum is 95 and 54, respectively. On average the participants identified 50 collocations and 28.5 distractors correctly. The median for correctly identified collocations is 41, while the median for correctly identified distractors is 21.5. The results scored by applying solution sheet one are shown in Table 7 and Figure 4 below.

	Total score	Collocations	Distractors
Minimum	36	23	10
Maximum	147	95	54
Average	78.5	50	28.5
Median	58.5	41	21.5

Table 7. General results WAT (s.sh. 1)

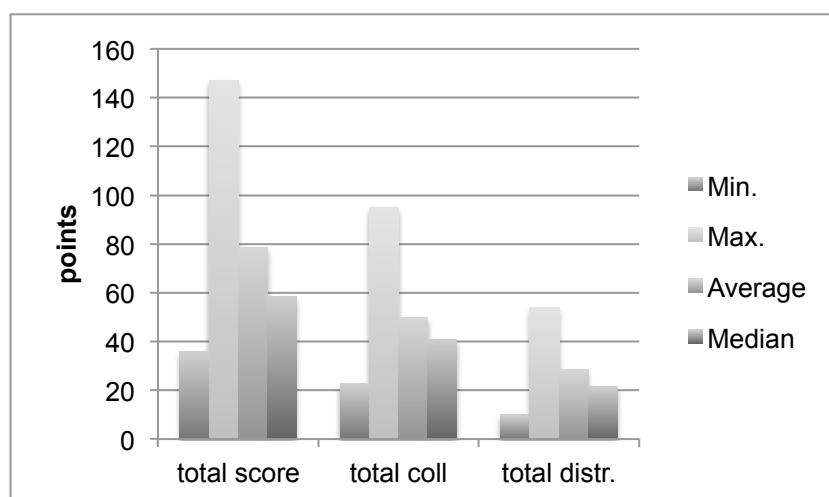


Figure 4. General results WAT (s.sh. 1)

Second, the results were also analysed with solution sheet two (i.e. every word combination appearing more than fifty times together in the CCL corpus is considered a collocation). The results changed drastically when applying this solution to the participants' tests, indicating that

word combination frequency does not necessarily play into whether a combination is considered as collocation or not. When applying this solution, the total number of possible collocations is 40 and the total number of distractors is 136, with the total possible score of 176 points staying unchanged. The minimum and maximum total score reached was 27 and 112, respectively, with a minimum of eleven identified collocations and 15 distractors and a maximum of 39 identified collocations and 82 distractors. When employing this solution the average total score of the participants is 64.8, with an average of 22 correctly identified collocations and 42.8 correctly identified distractors. The total score's median is 49, while the median for correctly identified collocations and distractors is 17 and 35.5 respectively, as shown by Table 8 and Figure 5 below.

	Total score	Collocations	Distractors
Minimum	27	11	15
Maximum	112	39	82
Average	64.8	22	42.8
Median	49	17	35.5

Table 8. General results WAT (s.sh. 2)

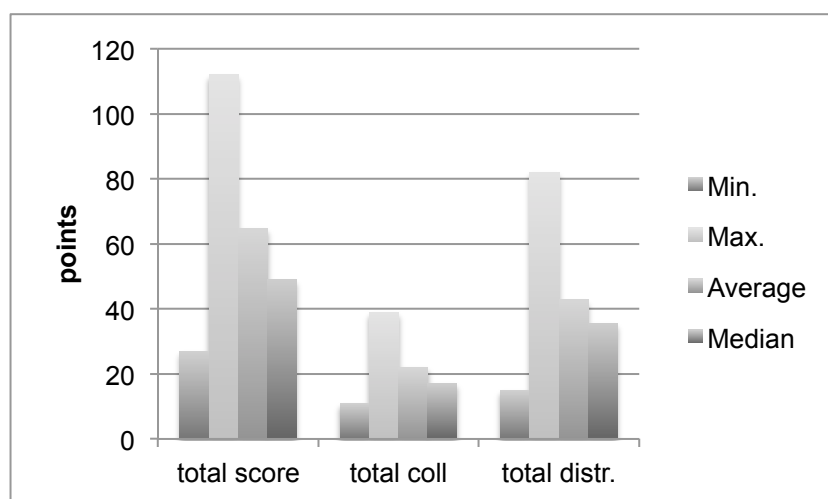


Figure 5. General results WAT (s.sh. 2)

In a third step, solution sheet three of Chen's (2016) study was also applied to the current study's tests. The possible number of collocations in this solution is 102 with 74 distractors and again a total of 176 points was possible. The minimum was 37 and the maximum score 153. On average the participants scored 82 points. The minimum and maximum of correctly identified collocations is 21 and 94 respectively, while the minimum and maximum of correctly identified distractors is 11 and 63, respectively. The participants identified an average of 49 collocations

and 33 distractors. The median total score is 56.5, while the median for correctly identified collocations and distractors is 38 and 22.5, respectively. The results are also shown in Table 9 and Figure 6 below.

	Total score	Collocations	Distractors
Minimum	37	21	11
Maximum	153	94	63
Average	82	49	33
Median	56.5	38	22.5

Table 9. General results WAT (s.sh. 3)

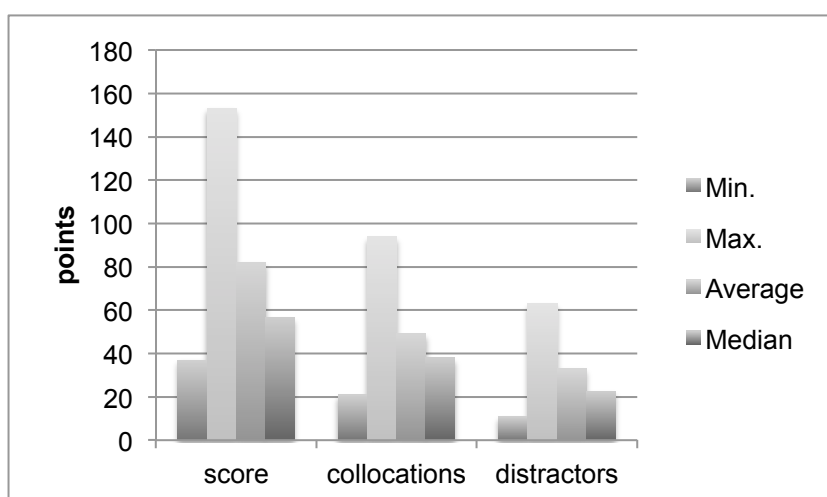


Figure 6. General results WAT (s.sh. 3)

To better illustrate the distribution of the participants' total scores, scatter plots of all three solution sheets (Figures 7, 8 and 9) are included below.

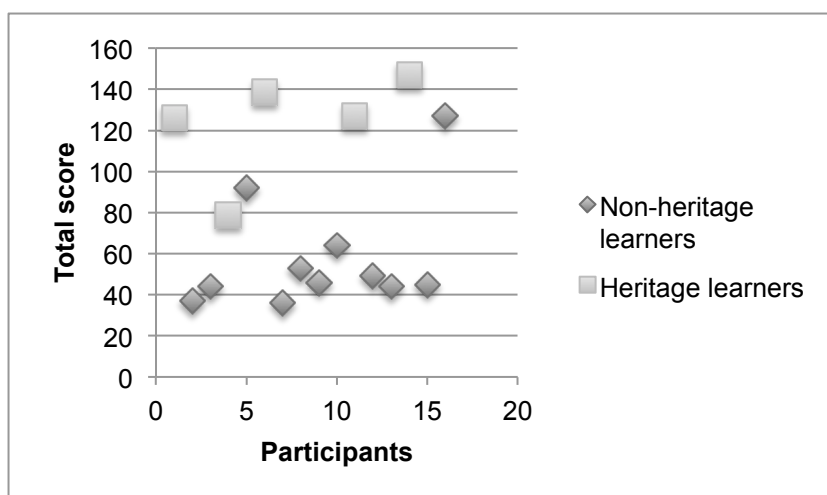


Figure 7. Scatter plot of total scores (s.sh. 1)

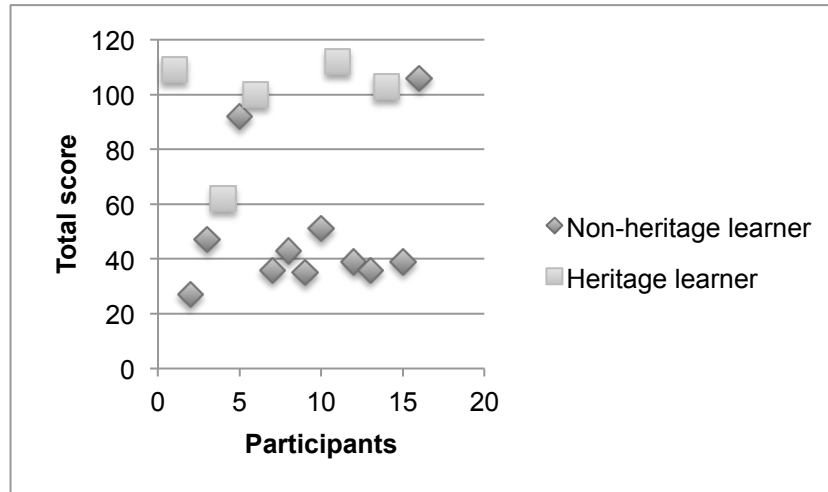


Figure 8. Scatter plot of total scores (s.sh. 2)

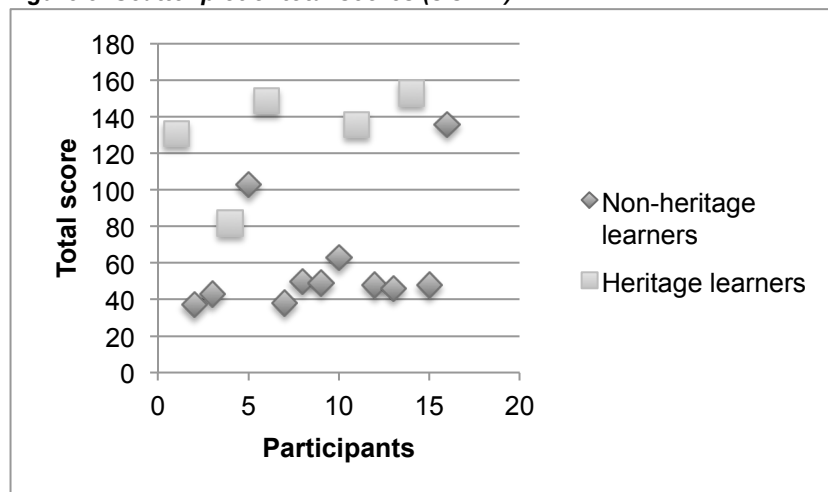


Figure 9. Scatter plot of total scores (s.sh. 3)

As mentioned in the section on data analysis in chapter six, the target words and word options that were marked as unknown by the participants are recorded but not included in the scoring of the test. However, as it is interesting to see how many words were unknown, these findings shall also be discussed here. Starting with the target words the minimum of unknown target words is zero and the maximum is 25, while on average 13.2 target words were unknown. The median for unknown target words is 16. Moving on to the unknown word options the minimum again is zero, while the maximum here is 22, which amounts to an average of 6.9 unknown word options and the median for unknown word options is 4.5. The results are shown in Table 10 and Figure 10 below.

	Unknown target words	Unknown word options	Blank items	Blank word options
Minimum	0	0	0	0
Maximum	25	22	18	4
Average	13.2	6.9	1.7	0.5
Median	16	4.5	0	0

Table 10. Unknown target words & word options, blank items & word options

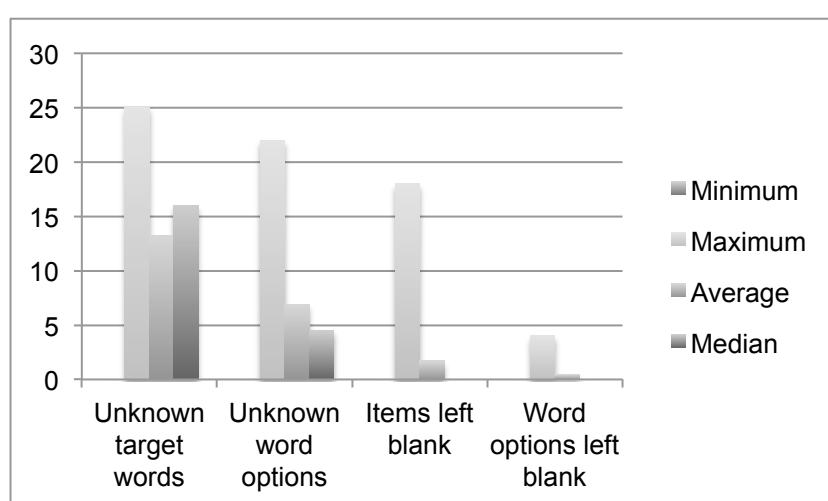


Figure 10. Unknown target words & word options, blank items & word options

What is of course at least as interesting as counting the number of unknown words is to look at the actual target words marked as unknown by the participants. Table 11 below shows, that the target word of item 33 充沛 [chōngpèi] (= ‘plentiful’, ‘vigourous’, ‘abundant’) was marked as unknown by twelve of the 16 participants, while the target word of item 31 舒畅 [shūchàng] (= ‘happy’, ‘entirely free from worry’) was marked as unknown by eleven participants. The third rank of the unknown target words is shared by four lexical items, all of them marked as unknown ten times. These target words are the words of items 26 锐利 [ruìlì] (= ‘sharp’, ‘penetrating’), 34 低廉 [dīlián] (= ‘cheap’, ‘inexpensive’, ‘low’), 40 肤浅 [fūqiǎn] (= ‘superficial’, ‘shallow’) and 44 凌乱 [língluàn] (= ‘in disorder’, ‘messy’).

Item number	target word	marked as unknown by ... participants
33	充沛 [chōngpèi] (= ‘plentiful’, ‘vigourous’,	12

	'abundant')	
31	舒畅 [shūchàng] (= 'happy', 'entirely free from worry')	11
26	锐利 [ruìlì] (= 'sharp', 'penetrating')	10
34	低廉 [dīlián] (= 'cheap', 'inexpensive', 'low')	10
40	肤浅 [fūqiǎn] (= 'superficial', 'shallow')	10
44	凌乱 [língluàn] (= 'in disorder', 'messy')	10

Table 11. Target words marked as unknown by the participants

When scoring the results it appeared that some of the participants left test items and/or word options completely blank (not marking any boxes, also not the unknown box). Accordingly these left blank items and word options were also recorded but were not included in the scoring. It should be noted that most participants did not leave any test items or word options blank. Therefore, the minimum for both test items and word options left blank is zero. The maximum of left blank test items is 18, while the maximum of left blank word options is four. Although only two participants left complete test items blank and three participants left some word options blank, an average for both was calculated: on average 1.7 test items and 0.5 word options were left blank. The results can also be seen in Table 8 and Figure 10 above. According to the study programme leader, Professor Dr. Weigelin-Schwiedrzik, all the words marked as unknown by the participants and presented in Table 11 above belong to a vocabulary that is rarely used during the course of the participants' studies at the Chinese department, as these words rather belong to a literary vocabulary. Again, this shows that choosing the vocabulary for such a test is key and may influence the results.

7.2.2. Identification of collocations in relation to word frequency

The words in the WAT can also be divided into three frequency levels: below 1000 (called high-frequency section in the following), between 1000 and 5000 (called mid-frequency section in the following) and above 5000 (called low-frequency section in the following). Thus, the present section focuses on the results in relation to word frequency.

Solution sheet 1

When applying solution sheet one (*counting as collocations only those which have a frequency of >10 in the CCL Corpus*), the total score possible for the high frequency words is 41 collocations and 15 distractors, making for a total of 56 points. The minimum scored for the high frequency section is 15 and the maximum is 50 points. The average score of the participants

was 30.9 points. The median of the total score is 26.5. The participants identified a minimum of nine and a maximum of 37 collocations correctly with an average of 22.5 and a median of 21. For the distractors of the high frequency group the minimum is two, the maximum is 15, the average is 8.4 and the median is 6.5.

The mid-frequency section contains 41 collocations and 19 distractors according to the first solution sheet. The minimum total score reached by the participants of the current study is eleven and the maximum is 45 with an average score of 27.6 and a median of 24. The participants identified a minimum of eight and a maximum of 30 collocations correctly, as well as a minimum of three and a maximum of 18 distractors. The average number of correctly identified collocations and distractors is 17.5 and 10.1, respectively, while the median for correctly identified collocations and distractors is 15.5 and 10, respectively.

Now consider the low-frequency section, the total score possible is 60 points with 29 collocations and 31 distractors. In this section the minimum of total score, as well as collocations and distractors is 0. The maximum score of this section is 52; the maximum number of correctly identified collocations is 28 and the correctly identified distractors 25. On average the participants scored a total of 19.8 and identified 10 collocations and 9.8 distractors correctly. The median for the total score is 12, while the median for correctly identified collocations and distractors is 6 and 5.5, respectively. Table 12 and Figure 11 below show the total scores of the participants across all three frequency levels.

High-frequency	Total score	Collocations	Distractors
Minimum	15	9	2
Maximum	50	37	15
Average	30.9	22.5	8.4
Median	26.5	21	6.5
Mid-frequency			
Minimum	11	8	3
Maximum	45	30	18
Average	27.6	17.5	10.1
Median	24	15.5	10
Low-frequency			

Minimum	0	0	0
Maximum	52	28	25
Average	19.8	10	9.8
Median	12	6	5.5

Table 12. Results WAT across frequency levels (s.sh. 1)

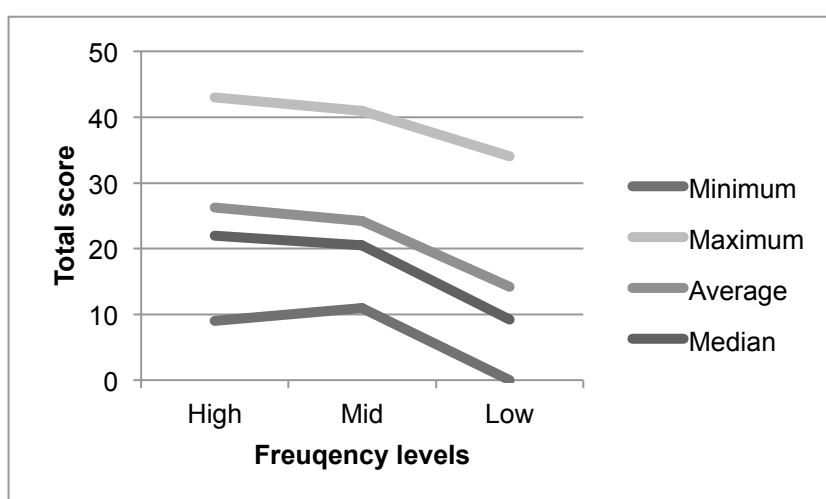


Figure 11. Results WAT across frequency levels (s.sh. 1)

Solution sheet 2

According to this solution sheet (counting as collocations only those which have a frequency of >50 in the CCL Corpus) the high-frequency words section consists of 25 collocations and 31 distractors, which amounts to a total of 56 points. The minimum and maximum of correctly identified collocations for this section are five and 24, respectively, while the minimum and maximum of correctly identified distractors is four and 23. The average of correctly identified collocations and distractors is 13.8 and 12.5, respectively. The minimum total score for the high-frequency section is nine points, while the maximum is 43. On average the participants scored 26.3 points.

Twelve collocations and 43 distractors make up mid-frequency words section, which amounts to 60 points. The average of the participants' total score on this section is 24.2, with a minimum of eleven and a maximum of 41. The participants identified an average of 7.1 collocations and 17.1 distractors correctly. The minimum and maximum of correctly identified collocations is three and twelve, respectively, while the minimum and maximum of correctly identified distractors is five and 32.

The low-frequency section contains three collocations and 57 distractors, thus containing a total of 60 points. The minimum of correctly identified collocations and distractors, as well as total score in this section is zero. The maximums are three collocations, 31 distractors and a total score of 34. On average the participants identified one collocation and 13.2 distractors correctly, while the average total score for this section is 14.2 points. The total scores of the participants across all three frequency levels are shown by Table 13 and Figure 12.

High-frequency	Total score	Collocations	Distractors
Minimum	9	5	4
Maximum	43	24	23
Average	26.3	13.8	12.5
Median	22	12.5	10.5
Mid-frequency			
Minimum	11	3	5
Maximum	41	12	32
Average	24.2	7.1	17.1
Median	20.5	6	15.5
Low-frequency			
Minimum	0	0	0
Maximum	34	3	31
Average	14.2	1.1	13.2
Median	9.5	0.5	9.5

Table 13. Results WAT across all frequency levels (s.sh. 2)

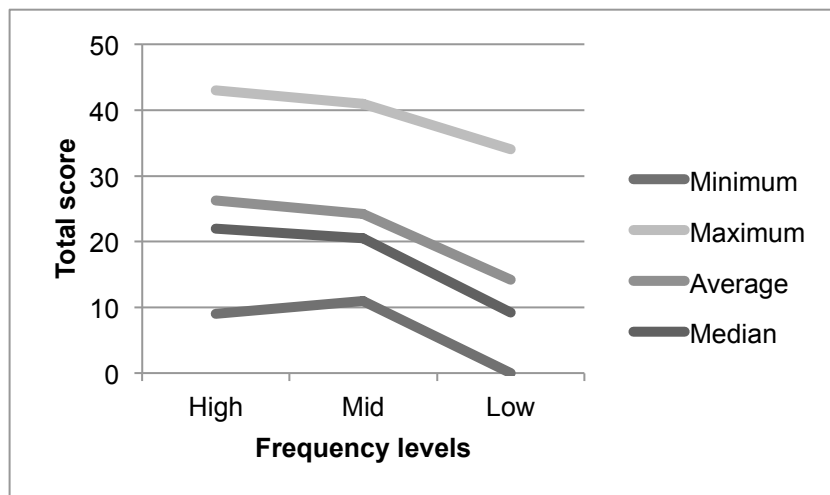


Figure 12. Results WAT across all frequency levels (s.sh. 2)

Solution sheet 3

Now we will look at the results in relation to word frequency by applying the solution developed by Chen (2016). We will start again with the high-frequency words section, which contains 37 possible collocations and 19 distractors (amounting to a total of 56 points). A minimum of nine, a maximum of 35 and an average of 22 collocations were identified correctly by the participants, while a minimum of four, a maximum of 19 and an average of 10.5 distractors were identified correctly. On average the total score of this section was 32.5 with a minimum of 17 points and a maximum of 50 points was reached by the participants.

The mid-frequency section contains 34 collocations and 26 distractors (amounting to a total of 60 points). The participants identified an average of 16.5 collocations and 12.8 distractors correctly. The minimum and maximum of correctly identified collocations is eight and 30, respectively. A minimum of three and a maximum of 24 distractors were identified correctly. The average total score of the mid-frequency words section amounts to 29.3, with a minimum of eleven and a maximum of 51.

Lastly, 31 collocations and 29 distractors (in total 60 points) make up the low-frequency section. The participants scored a minimum of 0 for collocations, distractors and total score. The maximum for all three is 29, 24 and 52, respectively. On average the participants reached a total score of 20.2, while the average for correctly identified collocations is 10.6 and the average for distractors is 9.6. The median for the total score is 13.5, while the median for correctly identified collocations and distractors is seven and 5.5, respectively. Table 14 shows the results across all three frequency levels, while Figure 13 below shows the participants' total scores across all three frequency levels.

High-frequency	Total score	Collocations	Distractors
Minimum	17	9	4
Maximum	50	35	19
Average	32.5	22	10.5
Median	27	20	8
Mid-frequency			
Minimum	11	8	3
Maximum	51	30	24
Average	29.3	16.5	12.8
Median	22.5	13	11
Low-frequency			
Minimum	0	0	0
Maximum	52	29	24
Average	20.2	10.6	9.6
Median	13.5	7	5.5

Table 14. Results WAT across all frequency levels (s.sh. 3)

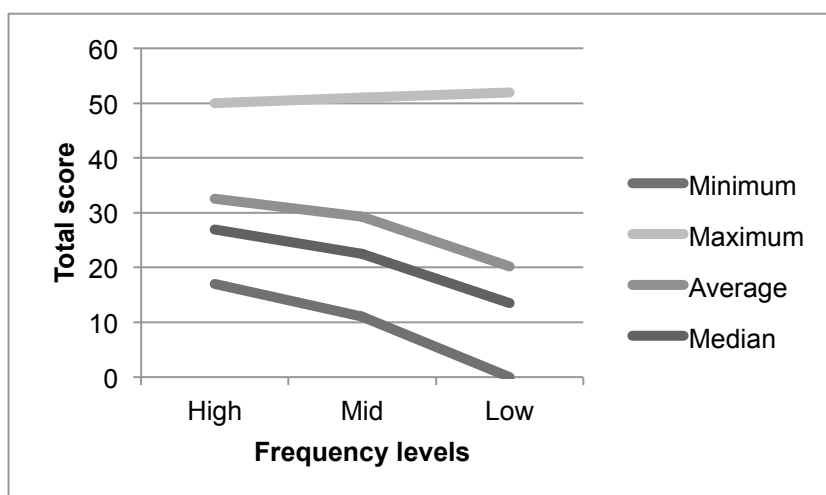


Figure 13. Results WAT across all frequency levels (s.sh. 3)

7.2.3. Differences between the two groups of participants

The participants of the current study can be categorized according to two relevant parameters: the first and third semester MA students on the one hand and heritage and non-heritage learners on the other hand. Consequently, the results will also be analysed in this respect.

First and third semester MA-students

First, the question whether there is a difference between the first and third semester MA students shall be considered and the results are therefore examined according to these two groups; again all three solution sheets will be considered.

Solution sheet 1

Starting with solution one, first semester MA students scored an average of 86.2 points as their total score, with a minimum of 37 and a maximum of 139. They identified an average of 53.6 collocations and 32.6 distractors correctly. The minimum and maximum for correctly identified collocations are 23 and 91, while the minimum and maximum for correctly identified distractors are 11 and 51. The total score's median is 85.5, while the median for correctly identified collocations and distractors is 53.6 and 32, respectively.

The third semester MA students on the other hand scored an average of 73.8 points with an average of 47.8 correctly identified collocations and 26 correctly identified distractors. They identified a minimum of 26 and a maximum of 95 collocations correctly. The minimum and maximum of correctly identified distractors are 10 and 54, respectively. The median of the total score is 51, while the median of correctly identified collocations and distractors is 37 and 16, respectively. Table 15 and Figure 14 show these results below.

First semester MA-students	Total score	Collocations	Distractors
Minimum	37	23	11
Maximum	139	91	51
Average	86.2	53.6	32.6
Median	85.5	53.5	32
Third semester MA-students			
Minimum	36	26	10

Maximum	147	95	54
Average	73.8	47.8	26
Median	51	37	16

Table 15. Results WAT: first vs. third semester MA-students (s.sh. 1)

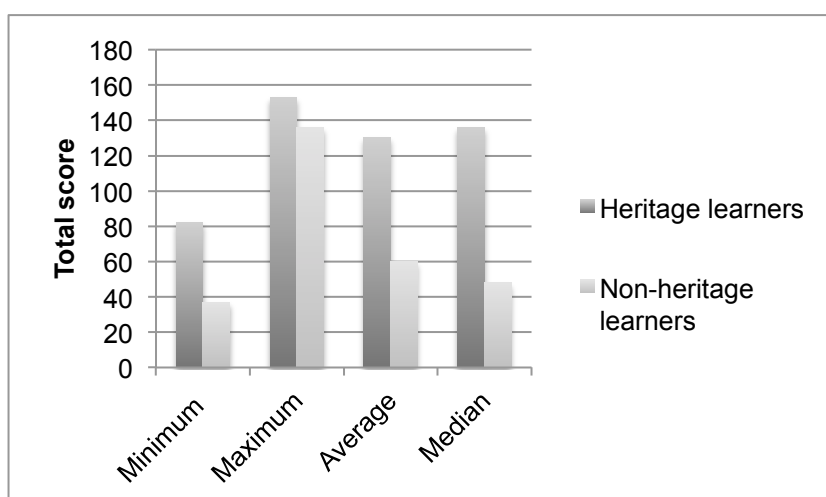


Figure 14. Results WAT: first vs. third semester MA-students (s.sh. 1)

Solution sheet 2

Moving on to solution two, first semester MA students scored an average of 72.8 points in total, identifying an average of 23.6 collocations and 49.2 distractors correctly. The minimum and maximum total score of the first semester MA students was 27 and 109. While the minimum of correctly identified collocations and distractors is eleven and 15, respectively, while the maximum of correctly identified collocations and distractors 36 and 75. The total score's median is 77, while the median of correctly identified collocations and distractors is 24.5 and 52.5, respectively.

The third semester MA students identified an average of 21 collocations and 39 distractors correctly, amounting to an average total score of 60 points. The minimum and maximum of the total score reached by these participants are 35 and 112 points. The minimum and maximum of correctly identified collocations is 14 and 39, while the minimum and maximum of correctly identified distractors is 18 and 82. The median of correctly identified collocations is 16, while the median for correctly identified distractors is 26. The total score's median is 41. Table 16 and Figure 15 show these results below.

First semester MA-students	Total score	Collocations	Distractors
Minimum	27	11	15
Maximum	109	36	75
Average	72.8	23.6	49.2
Median	77	24.5	52.5
Third semester MA-students			
Minimum	35	14	18
Maximum	112	39	82
Average	60	21	39
Median	41	16	26

Table 16. Results WAT: first vs. third semester MA-students (s.sh. 2)

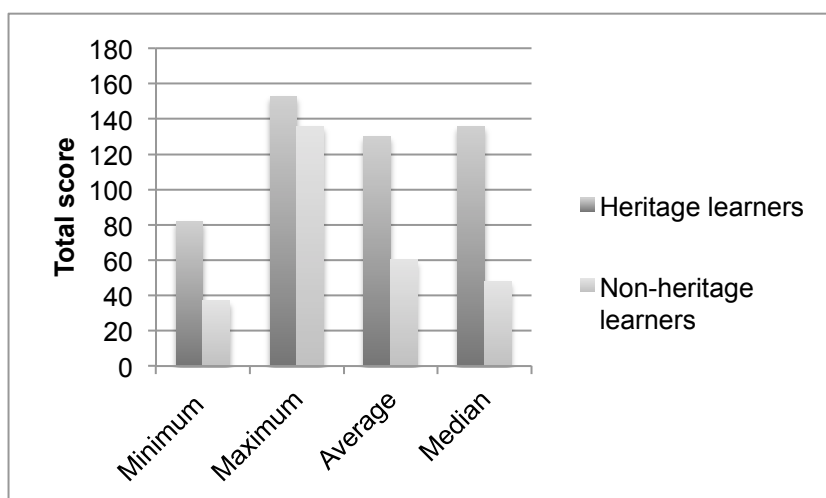


Figure 15. Results WAT: first vs. third semester MA-students (s.sh. 2)

Solution sheet 3

When applying Chen's solution to the participants' WAT, the results for these two groups of participants are the following. The first semester MA students identified a minimum of 21, a maximum of 90 and an average of 52.8 collocations correctly. The median for correctly identified collocations is 52.8. They identified a minimum of eleven, a maximum of 59 and an average of 38 distractors correctly, while the median for correctly identified distractors is 38. The total score reached on average by the first semester MA students amounts to 90.8 points, with a minimum of 37, a maximum of 149 points and a median of 92.5.

The third semester MA students scored an average total score of 76.7 points with a minimum of 38, a maximum of 153 and a median of 49.5 points. A minimum of 27 and a maximum of 94 collocations were identified correctly, while a minimum of 11 and a maximum of 63 distractors were identified correctly. On average the third semester MA students identified 46.8 and 29.9 collocations and distractors, respectively. The median for correctly identified collocations is 34.5, while the median for correctly identified distractors is 18.5. Table 17 and Figure 16 show these results below.

First semester MA-students	Total score	Collocations	Distractors
Minimum	37	21	11
Maximum	149	90	59
Average	90.8	52.8	38
Median	92.5	52.8	38
Third semester MA-students			
Minimum	38	27	11
Maximum	153	94	63
Average	76.7	46.8	29.9
Median	49.5	34.5	18.5

Table 17. Results WAT: first vs. third semester MA-students (s.sh. 3)

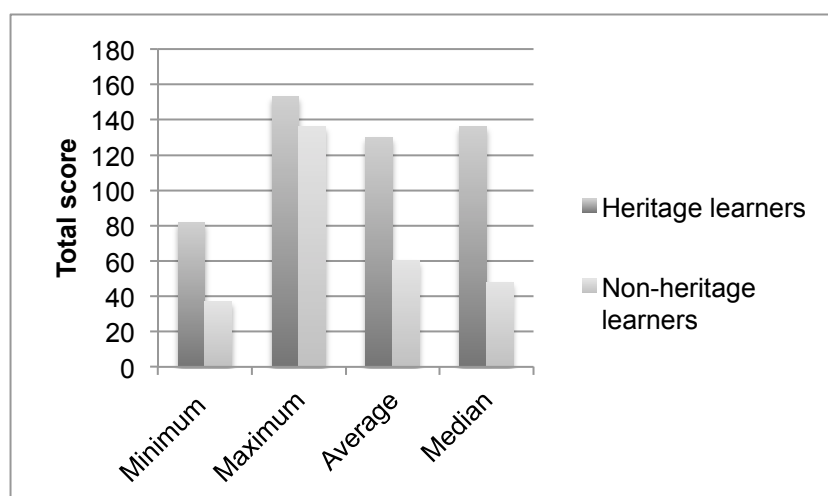


Figure 16. Results WAT: first vs. third semester MA-students (s.sh. 3)

Heritage and non-heritage learners of Chinese

Secondly, the participants can also be categorized into heritage and non-heritage learners of Chinese. Consequently it is also interesting whether there are differences between these two groups of learners concerning the correct identification of collocations. In total five out of 16 participants are heritage or native learners of Chinese. Again, we will look at the results scored with all three solutions.

Solution sheet 1

First, when scoring the WAT with solution one, the heritage learners scored an average total score of 123.6 points with a minimum of 79 and a maximum of 147 points. They identified a minimum of 52 and a maximum of 95 collocations and a minimum of 27 and a maximum 54 distractors correctly. On average they correctly identified 77.2 collocations and 46.6 distractors. The total score's median is 127 points, while the median of correctly identified collocations and distractors is 75 and 51, respectively.

The non-heritage learners identified an average of 37.6 collocations and 20.3 distractors correctly, with an average total score of 57.9 points. The minimum and maximum of correctly identified distractors are ten and 54. The minimum and maximum of the total score amounts to 36 and 127 points. The median of correctly identified collocations and distractors is 32 and 14, respectively, while the median of the total score is 46 points. These results are also shown in Table 18 and Figure 17 below.

Heritage learners	Total score	Collocations	Distractors
Minimum	79	52	27
Maximum	147	95	54
Average	123.6	77.2	46.6
Median	127	75	51
Non-heritage learners			
Minimum	36	23	10
Maximum	127	73	54
Average	57.9	37.6	20.3
Median	46	32	14

Table 18. Results WAT: heritage vs. non-heritage learners (s.sh. 1)

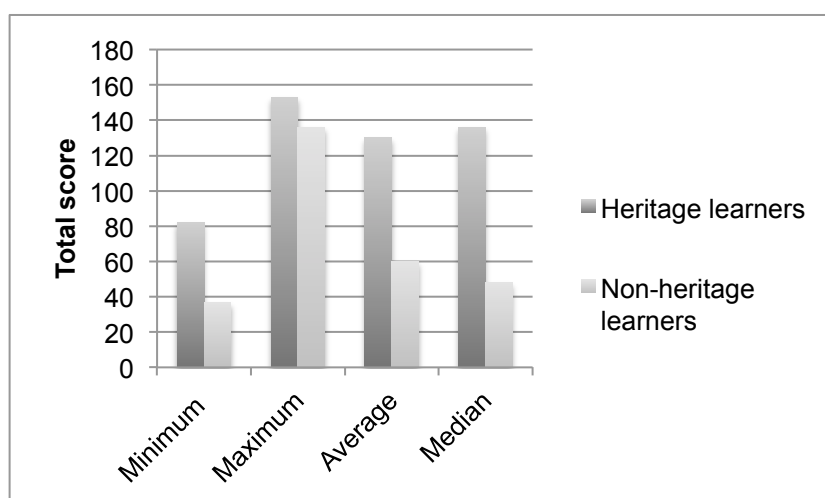


Figure 17. Results WAT: heritage vs. non-heritage learners (s.sh. 1)

Solution sheet 2

When applying solution sheet two, heritage learners scored a minimum of 62 and a maximum of 112 points and the average reached by them is 97.2. The median of the total score is 103 points. They identified a minimum of 19, a maximum of 39 and an average of 31.6 collocations correctly. As for distractors, they identified a minimum of 43, a maximum of 82 and an average of 65.6 correctly. The median for correctly identified collocations is 30, while the median for correctly identified distractors is 64.

The non-heritage learners reached an average total score of 50.1 with a minimum of 27, a maximum of 106 and a median of 39 points. A minimum of eleven and a maximum of 30 collocations were identified correctly, while a minimum of 15 and a maximum of 76 distractors were identified correctly. The average number of collocations identified correctly is 17.6, while the average number of correctly identified distractors is 32.4. The median for correctly identified collocations is 16, while the median for correctly identified distractors is 25. The results can also be found in Table 19 and Figure 18 below.

Heritage learners	Total score	Collocations	Distractors
Minimum	62	19	43
Maximum	112	39	82
Average	97.2	31.6	65.6
Median	103	30	64
Non-heritage learners	Total score	Collocations	Distractors
Minimum	27	11	15
Maximum	106	30	76
Average	50.1	17.6	32.4
Median	39	16	25

Table 19. Results WAT: heritage vs. non-heritage learners (s.sh 2)

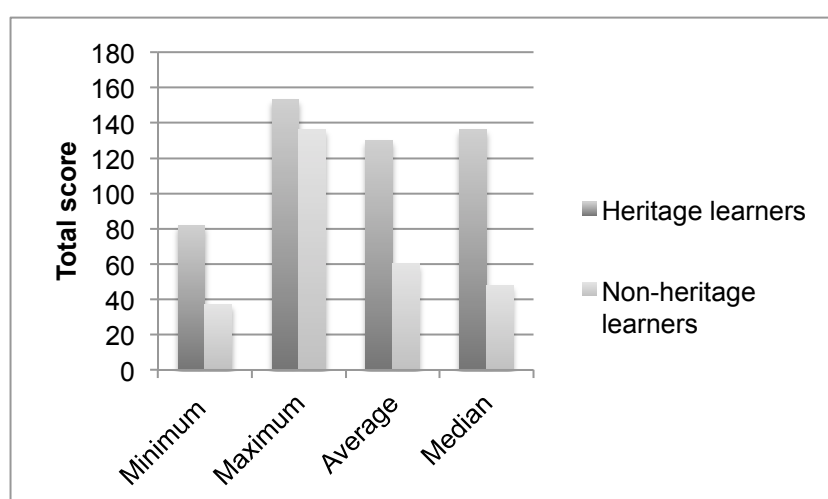


Figure 18. Results WAT: heritage vs. non-heritage learners (s.sh. 2)

Solution sheet 3

Lastly, the following are the results when the WAT is scored with Chen's solutions (i.e. solution sheet three). An average total score of 130.2 was reached by the heritage learners with a minimum of 82, a maximum of 153 and a median of 136 points. They identified an average of 63.5 collocations and 54 distractors correctly. The minimum and maximum number of correctly identified collocations are 52 and 94, respectively, while the minimum and maximum number of distractors are 30 and 63, respectively. The median of correctly identified collocations is 73, while the median for correctly identified distractors is 59.

The non-heritage learners on the other hand reached an average total score of 60.1 points. The minimum total score is 37 and the maximum total score is 136. The median of the total score is 48. They identified an average of 36.7 collocations and 23.3 distractors correctly. The minimum number of collocations and distractors identified correctly are 21 and eleven, respectively, while the maximum number of collocations and distractors identified correctly are 73 and 63, respectively. The median for correctly identified collocations is 33, while the median for correctly identified distractors is 18. Again, these results are shown by Table 20 and Figure 19 below.

Heritage learners	Total score	Collocations	Distractors
Minimum	82	52	30
Maximum	153	94	63
Average	130.2	63.5	54
Median	136	73	59
Non-heritage learners			
Minimum	37	21	11
Maximum	136	73	63
Average	60.1	36.7	23.3
Median	48	33	18

Table 20. Results WAT: heritage vs. non-heritage learners (s.sh. 3)

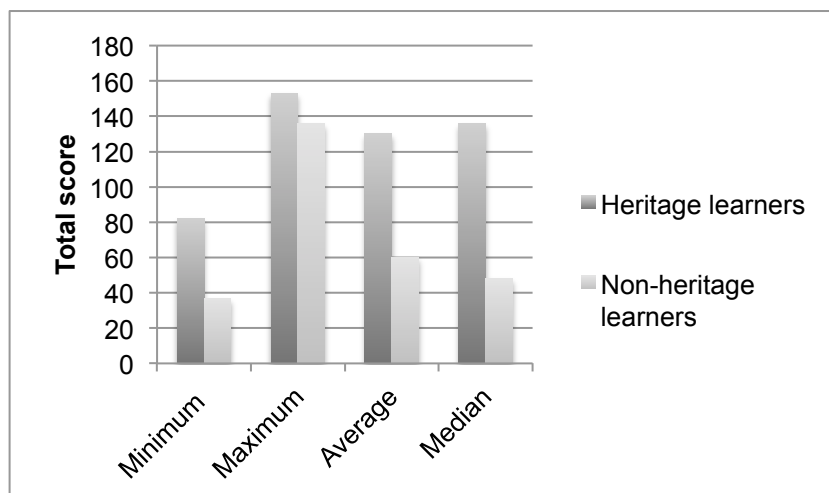


Figure 19. Results WAT: heritage vs. non-heritage learners (s.sh. 3)

8. Discussion and Conclusion

This chapter will discuss the state of the participants' collocational knowledge and how it seems to be impacted by word frequency level based on the results presented in the previous chapter. Whether other factors (e.g. language learning background, stay abroad in China) play into their collocational knowledge will also be discussed. The results of the current study will also be compared to the results of its model study conducted by Chen (2016). After discussing the results, implications for teaching and learning CFL will be explored.

8.1. Discussion of the results

First our focus will turn to overall state of the participants' vocabulary knowledge in the context of the current study. Although this study is primarily interested in the participants' depth of vocabulary knowledge, a conclusion about the participants' vocabulary breadth can be drawn from the number of known target words as well. Out of a total of 44 target words the participants knew an average of 30.8 and a median of 28 target words, which amounts to 70% and 63.6%, respectively. In other words, the participants knew about two thirds of the target words. As the target words were selected from three different frequency levels (i.e. <1.000, 1.000-5.000, >5.000) so that one third of the words actually belonged to the low frequency band, the participants' vocabulary size seems to be quite large, covering also the mid-frequency vocabulary range. In contrast, the accuracy rate for correctly identified collocations is lower than the accuracy rate of the total score, as can be seen in Table 21 below. This suggests a

disbalance in the development of the participants' depth of vocabulary knowledge and their vocabulary size: while they had acquired a large number of words, the depth or "quality" of their knowledge regarding a large portion of these words still needs improvement. Evidence from earlier studies (e.g. Lennon, 1996:35) suggests that the vocabulary knowledge of advanced L2 learners may be "hazy" when understanding ambiguities of meaning and "collocational restrictions", despite them having a large vocabulary size.

Note: Average accuracy rate / Median accuracy rate

	Average accuracy rate	Median accuracy rate
Solution sheet 1	45%	36.9%
Solution sheet 2	50%	38%
Solution sheet 3	48%	37%

Table 21. Average and median accuracy rate of the participants' total score

Now consider the results in relation to word frequency and whether it impacts the participants' collocational knowledge. According to research (e.g. Greidanus, Beks & Wakely, 2005) the word knowledge of advanced L2 learners is deeper regarding "frequent, earlier acquired words" than it is regarding "less frequent, more recently acquired words." As shown by Table 22 below, this pattern can also be found in the results of the current study, since accuracy dropped as word frequency decreased:

	Accuracy rates: High-frequency	Accuracy rates: Mid-frequency	Accuracy rates: Low-frequency
Solution sheet 1	55.2% / 47.3%	46% / 40%	33% / 20%
Solution sheet 2	46.9% / 39.3%	40.3% / 34.2%	23.6% / 15.8%
Solution sheet 3	58% / 48.2%	48.8% / 37.5%	33.6% / 22.5%

Table 22. Accuracy rate by frequency of target word

The more dramatic drop of the participants' accuracy rate is clearly from the mid-frequency level to the low-frequency level, as can be seen in Table 20 above. This shows that the results of these participants follow the pattern shown by earlier research mentioned above.

One of the research questions of this thesis (RQ3) is also whether other factors (i.e. which of the independent variables) have an impact on the participants' collocational knowledge. If we consider the scatter plots from the previous chapter (i.e. Figures 7, 8 and 9 on pages 92-93), it

becomes apparent that the results of the majority of participants are within a rather narrow range of points. This suggests that most factors included in the demographic questionnaire (e.g. time spent learning CFL, stay abroad in China, etc.) do not greatly impact the participants' collocational knowledge, the one exception being that of L1 (i.e. whether the participants grew up speaking Chinese with their family members). The results presented in the previous section on the scores of the heritage and non-heritage learners show a huge difference between these groups of learners, as their total scores (shown by Figures 17, 18 and 19 in the previous chapter on pages 106, 107, 109, respectively) and also their average and median accuracy rates differ strongly, as shown by Table 23 below. This indicates that the L1 of a person impacts a person's collocational knowledge strongly, as was to be expected.¹³

Note: Average accuracy rate / Median accuracy rate

	Accuracy rates: heritage learners	Accuracy rates: non-heritage learners
Solution sheet 1	70.2% / 72.1%	32.9% / 26.1%
Solution sheet 2	55.2% / 58.5%	28.5% / 22.1%
Solution sheet 3	73.9% / 77.3%	34.1% / 27.3%

Table 23. Accuracy rate of heritage and non-heritage learners

Three participants are so-called outliers of these ranges, which is why a closer look will be taken at the background of these participants in order to perhaps find factors influencing the results of these three test-takers. The first outlier to be discussed is from the heritage learner group. Compared to the other members of this group this participant (called P4 in data analysis) scored a relatively low total score (i.e. 79 points with solution sheet one, 62 points with solution sheet two and 82 points with solution sheet three), while all other member of this group scored above 100 points with all solution sheets. This participant noted Chinese as her only L1, but when looking at the information provided by her demographic background questionnaire, she

¹³ However, it was noted by Professor Dr. Weigelin-Schwiedrzik that the majority of heritage learners come from a non-academic background, speak various dialects at home and in most cases could not read or write Chinese before starting their studies at the Chinese Department (except for the students who previously visited a Chinese weekend school). This shows that they indeed make big progress during their studies while learning Chinese as an educational language.

also noted not to have stayed abroad in a country with Chinese as its native language so far. This, then, seems to be a Chinese heritage speaker, who probably speak Chinese as a family language but has only very loose ties to the wider Chinese language community. The other two outliers (called P5 and P16 in data analysis) belong to the non-heritage learner group. The first of these two outliers (P5) obtained slightly higher scores than P4, the outlier of the heritage learner group, and lay thus at least 30 points above the other members of the non-heritage learner group with any of the three solution sheets (i.e. 92 points with solution sheet one, 100 points with solution sheet two and 103 points with solution sheet three). When looking at this participant's answers to the demographic background questionnaire, two factors suggest themselves as potentially impacting this participant's results: first, although this participant did not list Chinese as his L1, he stated that he had family members who speak Chinese as their native language and he "sometimes" speaks Chinese with them. The second factor that could also have an effect is that he spent ten months in a Chinese speaking country to study abroad. It thus seems that P4 and P5 might in fact be of a similar language background, with Chinese as heritage language but strongly dominant in the language of the environment (presumably German), even though P5 did not self-classify as heritage speaker. The advantage of P5 over P4 appears to stem from P5's study abroad experience in a Chinese-speaking country. The second outlier of the non-heritage learner group (P16) scored even higher – above 105 points with any of the three solution sheets (i.e. 127 points with solution sheet one, 106 points with solution sheet two and 136 points with solution sheet three). Again, two factors appear to have an impact on this participant's results. First, this participant stated that he attended a Chinese school on Saturdays for two and a half years. Which in fact suggests that he/she too has some Chinese heritage in her family background without self-classifying as Chinese heritage speaker in the demographic questionnaire. Secondly, he spent two years studying abroad in a Chinese speaking country. Thus, showing that immersing oneself in the language as much as possible seems to have a strong effect on collocational knowledge, but also, by implication, on general language proficiency. Contrary to expectations, the demographic background questionnaire indicated that length of study of CFL was not a meaningful indicator of depth of vocabulary knowledge. However, it should be noted again that the sample is too small to be representative.

The results of the WAT also show that the first semester MA students scored higher than the third semester MA students, which is rather unexpected. On closer examination of the demographic background questionnaire the reason for this becomes apparent. The first semester MA group consists of six participants, three of which are heritage learners and thus reached high scores. The third semester MA group on the other hand consists of ten

participants of which only two are heritage learners. It would be interesting to compare these two groups excluding the heritage learners. However, the groups would be even smaller, as then the group of first semester MA students would only consist of three participants, while the group of third semester MA students would only consist of eight participants. Although this leads to the results not being meaningful, it is still interesting to compare the first and third semester non-heritage MA students in order to see whether there are results of non-heritage learners that stand out in comparison to others. As discussed above, two participants of the non-heritage learner groups stand out, one in the first semester MA group (i.e. P5) and one in the third semester group (i.e. P16). However, the results of the other participants of both groups are too similar to be able to see a trend of improvement. To clarify this, the points scored by the non-heritage learners of both semesters (according to solution sheet one as an example) have been listed in Table 24 below. The two clear outliers (P5, P16) have been underlined in the table.

Non-heritage first semester participants	Total scored points (s.sh.1)
P2	37
P3	44
<u>P5</u>	<u>92</u>
Non-heritage third semester participants	Total scored points (s.sh.1)
P7	36
P8	53
P9	46
P10	64
P12	49
P13	44
P15	45
<u>P16</u>	<u>127</u>

Table 24. Results of non-heritage learners (s.sh.1)

Table 24 above clearly shows that most other results all lie in a similar area, except for maybe P8 and P10 who scored slightly higher. Taking a look at their answers in the demographic questionnaire, reasons for this slightly higher score may be found. P8 stated that he has been learning CFL for four to five years and spent three months abroad in a Chinese speaking

country. P10 on the other hand has been learning CFL for less than two years but spent one year abroad in a Chinese speaking country (living and attending school).

Comparing the results of the current study with its model study by Chen (2016) shows that on average the participants of this current study were able to identify a slightly higher number of target words than the participants of Chen's study, as they identified an average of 30.8 target words, while the participants of Chen's study identified an average of 28 target words. The average accuracy rate of the participants in the current study is also higher than that of Chen's participants: 45% to 50% (depending on the solution sheet) compared to 37% in Chen's study. This shows that the participants of the current study were able to identify more of the word options. Another difference in the results can be found in the effect that word frequency had on the accuracy rate and thus the participants' ability to identify collocations. While the current study's participants' accuracy rate had a more dramatic drop from the mid- to the low-frequency word section, the accuracy rate of Chen's study's participants dropped more dramatically from the high- to the mid-frequency section. This indicates that the participants of the current study know a large number of words in both the high- and mid-frequency section. Additionally, the participants of the current study scored higher than the participants of Chen's study in the low-frequency section as well: while Chen's study indicated an average accuracy rate of 21% for the low-frequency section, the current study's participants had an accuracy rate of around 33% depending on the solution sheet. The word knowledge of the participants of Chen's study thus seems to be distributed differently, as the results indicate that while they know a large number of high-frequency words, their knowledge is more limited when it comes to mid- and low-frequency words. The fact that the participants of the current study know more mid- and low-frequency words than the participants of Chen's study may be due to the different designs in curricula and thus a difference in what vocabulary is focused on in the Chinese language courses the respective participants attended. As mentioned in chapter six on research methodology, the Viennese Chinese Studies' curriculum is more focused on academic language, which may be why the participants' know more words also in the mid-frequency level.

8.2. Implications for teaching and learning CFL

According to researchers one central characteristic of advanced L2 proficiency is the knowledge of the "exact meaning and usage" of words (Kubler, 2002:111), which is also called the "precision of lexicon" (Leaver & Shekhtman, 2002:26). This is also reflected by the results of the current study, as the participants knew a large quantity of target words but their accuracy in identifying collocations lagged behind the knowledge of target words. Because her study

indicated a similar pattern, Chen (2016) suggested that a possible reason for this pattern “could be a lack of emphasis in classroom instruction on establishing and continually broadening the learners’ vocabulary networks” (Chen, 2016:126). This may be due to the fact that words are not taught with emphasis on the relationship(s) between words but are still taught as individual units. Thus, a recommendation for CFL curriculum design can be drawn from these results: systematic and explicit instruction and practice on the connections among words should be included in language teaching so as to sensitize learners to these connections.

Considering the results of the mid-frequency words section again, we can draw the conclusion that the focus of the Viennese language course curriculum on a more academic and professional vocabulary definitely has an effect on the participants’ word knowledge in this section, as they scored much higher than Chen’s participants. However, the current study’s participants’ knowledge of low-frequency words should still be expanded.¹⁴ This could be achieved through extensive reading of texts of various genres and styles. As discussed in chapter four on learning vocabulary, extensive reading is also helpful to acquire and retain word meanings. Varied language contexts are particularly helpful for learners of higher language proficiency to learn word meaning and use implicitly and to be able to infer word meaning from context rather than relying on a dictionary. This indicates the importance of extensive reading also in a CFL context.

Furthermore, studying abroad offers a rich environment for vocabulary learning, which is why one hypothesis of the present researcher was that whether a participant had studied abroad could have a clear effect on his or her ability to identify collocations. However, apart from two participants in the non-heritage learner group, this seemed to be a wrong assumption because 14 out of 16 participants stated they had stayed or studied abroad in a country with Chinese as its native language (seven participants stayed between three to six months, four participants between seven to twelve months and one participant each stayed two years, six years and stated “not countable”, respectively). This might be due to the fact that according to literature (e.g. Llanes, 2010) studying abroad may primarily influence learners’ spoken language proficiency rather than reading or writing. This was also indicated by the results of Chen’s (2016:129) interviews. As improving spoken language often is the main motivation for language

¹⁴ However, it should be noted again that the focus concerning vocabulary in the master’s programme at the Chinese Department is focused on academic language in order to prepare the students for the use of Chinese sources for their master’s theses. Thus, the low-frequency words belonging to literary vocabulary is not necessary in order to reach this goal.

learners to study abroad (e.g. Davidson, 2010), their focus thus naturally lies on spoken or colloquial language (口语 [kǒuyǔ]) not on written language (书面语 [shūmiànyǔ]), which is why more academic, professional or literary vocabulary is often neglected during stays abroad. When considering the information given in the demographic background questionnaire of the current study, the answers of the third outlier discussed above (i.e. P16) further show that he not only studied abroad much longer than the other participants, but additionally indicated that not only did he take Chinese language courses but also took academic content courses (e.g. on Taiwanese history) which were held in Chinese. This shows that a carefully planned and structured stay abroad can have a very positive effect on vocabulary and reading development.

8.3. Limitations

Although this study gained interesting results with the help of the Word Associates Test (WAT) format, it also has its limitations. First, the sample size of sixteen participants is rather small due to the fact that the MA programme at the Chinese department is also rather small.¹⁵ But in order to gain a larger sample size, the test would have to be repeated over a longer period of time in order to capture more students as they complete their language requirement. Due to the small sample size, the results of the study may not be representative of all MA students at the Vienna Chinese department, let alone CFL learners as a whole.

A second limitation would be the fact that the WAT only included one word class, namely adjectives, and only those comprised of two characters. Thus, the findings are also not necessarily representative of words belonging to other word classes or of other composition structures. Additionally, the number of included target words (i.e. 44 target words were included in the scoring of the WAT with approximately 15 target words for each frequency level) is also not representative of the entire Chinese vocabulary.

Additionally, at this point the question whether the model study has proved worthwhile should be considered. The selection of vocabulary plays an important role here, as vocabulary is central to such a test as WAT. In her thesis, Chen (2016:55-57) listed the criteria she used for selecting the words for the test items of her WAT, which were also discussed in section 6.1. *Model study*. One problem arising with these selection criteria is that the factor of language

¹⁵ In the term the current study was conducted in (i.e. winter term 2019/20) the MA programme of the Chinese department had a total of 74 students, with six of those being first semester MA students. (Studienservice und Lehrwesen, 2020)

proficiency was not considered (in form of which HSK levels the words belong to, which was also discussed in section 6.2. *Adapted test*). This may have lead to words being included in the test that are rarely learned by CFL learners.

Another limitation may arise due to the fact that Chen (2016) chose collocations of adjectives and nouns, as she did not provide an explanation as to why she chose this combination for the test and not other combinations (such as verbs and objects, which are also important possible combinations). Consequently another problem arises, as it is not necessarily true that adjectives are representative of the Chinese language as a whole.

Furthermore, the scoring methods show that solution sheet one, which considers every combination appearing in the CCL corpus and Chen's (2016) scoring method (i.e. solution sheet three), which was composed with the help of a collocation dictionary, achieve similar results. However, solution sheet two, which only considers combinations as collocations if they appear more than 50 times in the CCL corpus, achieves different results. This shows that the method of working out the scoring sheets may be an additional factor impacting the test's results.

8.4. Conclusion

The current study was conducted because of a need to better understand the vocabulary development of CFL learners. This is true for all proficiency levels, as well as for both breadth and depth of vocabulary knowledge. Because one study cannot include a test for all aspects of vocabulary knowledge, the focus on depth of vocabulary knowledge, specifically on collocations, was chosen. Due to the inclusion of low-frequency words in the Word Associates Test (WAT) the study was conducted with MA students who already have a higher proficiency in the Chinese language. The results of the WAT have shown that the participants know a large quantity of target words and were also able to identify a large number of collocations correctly. However, differences were evident relating to word frequency: the lower the word frequency, the fewer collocations were identified correctly. What was also clearly discernible was that the focus on academic and professional vocabulary in the curriculum at the Chinese department at the University of Vienna seems to have a positive effect on the participants' knowledge of mid- and low-frequency words.

Although this study may help to further close the research gap, further research in this area, as well as other areas of vocabulary knowledge is important due to the limitations of this study. Hence research should be continued and an investigation of vocabulary depth including other

word classes and also other dimensions of word associations is of interest to further understand the development of vocabulary knowledge in CFL learners. As is widely acknowledged among researchers and language testing experts, the vocabulary knowledge of learners is an excellent indicator of general language proficiency, which is why a comparison of learners with different CFL proficiency levels should also be explored.

Bibliography

- Aitchison, J. (2012) *Words in the mind. An introduction to the mental lexicon*. 4th edition. New Jersey, USA, John Wiley & Sons, Inc.
- Bassetti, B. (2005) Effects of writing systems on second language awareness: word awareness in English learners of Chinese as a foreign language. In: *Second language writing systems*. Clevedon, UK, Multilingual Matters, pp. 335-356.
- Beglar, D. & Nation, P. (2014) Assessing vocabulary. In: Kunnan, A.J. (ed.) *The companion to language assessment*. Volume I: Abilities, contexts, and learners. New Jersey, USA, Wiley Blackwell, pp.172-184.
- Benson, M., Benson, E. & Ilson, R. (2009) *The BBI combinatory dictionary of English. A guide to word combinations*. 3rd edition. Amsterdam [u.a.], John Benjamins Pub. Co.
- Boggards, P. & Laufer-Dvorkin, B. (2004) *Vocabulary in a second language: selection, acquisition, and testing*. Amsterdam [u.a.], John Benjamins Pub. Co.
- Chen, T. (2016) *Exploring depth of vocabulary knowledge among CFL learners of high proficiency*. Iowa, USA, University of Iowa.
- Chen, X. & Pasquarella, A. (2017) Learning to read Chinese. In: Verhoeven, L. & Perfetti, C. (eds.) *Learning to read across languages and writing systems*. Cambridge [u.a.]: Cambridge University Press, pp. 31-56.
- Coleman, J.A. (2010) Study abroad and SLA: defining goals and variables. In: Berndt, A. & Kleppin, K. (eds.) *Sprachlehrforschung: Theorie und Empirie*. Frankfurt a. M. [u.a.], Lang, pp.181-196.
- Cronbach, L. (1942) An analysis of techniques for diagnostic vocabulary testing. *The Journal of Educational Research*. 36 (3), 206-217.
- Daller, H., Milton, J. & Treffers-Daller, J. (eds.) (2007) *Modelling and assessing vocabulary knowledge*. Cambridge [u.a.], Cambridge University Press.
- Dewaele, J., Witney, J., Saito, K. & Dewaele, L. (2017) Foreign language enjoyment and anxiety: the effect of teacher and learner variables. *Language Teaching Research*. 22 (6), 676-697.
- Dörnyei, Z. (2007) *Research methods in applied linguistics*. Oxford, Oxford University Press.
- Ellis, R. & Barkhuizen, G.P. (2005) Collecting samples of learner language. In: Ellis, R. & Barkhuizen, G. (eds.) *Analysing learner language*. Oxford, Oxford University Press, pp. 15-50.

- Everson, M. (1998) Word recognition among learners of Chinese as a foreign language: investigation the relationship between naming and knowing. *Modern Language Journal*. 82 (2), 194-204.
- Everson, M. & Ke, C. (1997) An inquiry into reading strategies of intermediate and advanced learners of Chinese as a foreign language. *Journal of the Chinese Language Teachers Association*. 32 (1), 1-20.
- Gärdenfors, P. (2014) A semantic theory of word classes. *Croatian Journal of Philosophy*. 41, 179-194.
- Garman, M. (1990) Accessing the mental lexicon. In: Garman, M. (ed.) *Psycholinguistics*. Cambridge [u.a.], Cambridge University Press, pp. 239-300.
- Guder-Manitius, A. (2009) Zur chinesischen Sprache und Schrift. In: Guder-Manitius, A. (ed.) *Sinographemdidaktik. Aspekte einer systematischen Vermittlung der chinesischen Schrift im Unterricht Chinesisch als Fremdsprache*. Tübingen, Julius Groos Verlag, pp. 6-66.
- Guo Rui [郭锐] (2002) 现代汉语词类研究 [Xiàndài hànyǔ cílèi yánjiū] *Research on modern Chinese word classes*. Shanghai, China: 商务印书馆 [Shāngwù yìn shūguǎn] The Commercial Press.
- Hu Renfen [胡韧奋] & Xiao Hang [肖航] (2019) 面向二语教学的汉语搭配知识库构建及其应用研究 [Miànxàng èryǔ jiàoxué de hànyǔ dāpèi zhīshìkù gòujiàn jí qí yìngyòng yánjiū] *Research on the construction and application of Chinese collocation knowledge base for second language teaching*. 语言文字应用 [Yǔyán wénzì yìngyòng] *Applied Linguistics*. (1), 135-144.
- Hu Yushu [胡裕树] & Fan Xiao [范晓] (1994) 动词形容词的“名物化”和“名词化” [Dòngcí xíngróngcí de ‘míngwùhuà’ hé ‘míngcíhuà’] The ‘nominalisation’ and ‘nounisation’ of verbs and adjectives. 中国语文 [Zhōngguó yǔwén] *Chinese language*. (2), 81-85.
- Ježek, E. (2016) *The lexicon. An introduction*. Oxford textbooks in linguistics. Oxford, Oxford University Press.
- Laufer, B. (2003) Vocabulary acquisition in a second language: do learners really acquire most vocabulary by reading? Some empirical evidence. *The Canadian Modern Language Review*. 59 (4), 567-587.
- Laufer, B. & Goldstein, Z. (2004) Testing vocabulary knowledge: size, strength and computer adaptiveness. *Language Learning*. 54 (3), 399-436.
- Laufer, B. & Nation, P. (1995) Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics*. 16 (3), 307-322.
- Lee-Thompson, L. (2008) An investigation of reading strategies applied by American learners of Chinese as a foreign language. *Foreign Language Annals*. 41 (4), 702-721.

- Li Jun [李军] (2008) 词语搭配问题与语言研究 [Cíyǔ dāpèi wèntí yǔ yǔyán yánjiū] Word collocation and linguistics. *中华文化论坛 [Zhōnghuá wénhuà lùntán] Chinese Culture Forum*. (4), 73-78.
- Li Rulong [李如龙] (2011) 论汉语词汇的多元系统 [Lùn Hànyǔ cíhuì de duōyuán xìtǒng] On the polysystem of Chinese vocabulary. *江苏大学学报 (社会科学版) [Jiāngsū dàxué xuébào (shèhuì kēxué bǎn)] Journal of Jiangsu University (Social science edition)*. 1, 48-54.
- Liao Guangrong [廖光蓉] (2019) 汉语词类超常活用及典型性探讨 [Hànyǔ cílèi chāocháng huóyòng jí diǎnxíngxìng tàntǎo] An exploration of the extraordinary usage and typicality of Chinese word classes. *湖南大学学报 (社会科学版) [Húnán dàxué xuébào (shèhuì kēxuébǎn)] Journal of Hunan University (Social Science Edition)*. 33 (1), 105-112.
- Liu, D. (2018) Literature review of the breadth and depth of vocabulary. *Journal of language teaching and research*. 9 (5), 1002-1008.
- Liu Rushan [刘汝山] (2001) 汉语词语搭配类型及其在《汉英搭配词典》中的收录问题 [Hànyǔ cíyǔ dāpèi lèixíng jí qí zài 'Hàn yīng dāpèi cídiǎn' zhōng de shōulù wèntí] Collocation types of Chinese words and their inclusion in the 'Chinese-English collocation dictionary'. *青岛海洋大学学报 (社会科学版) [Qīngdǎo hǎiyáng dàxué xuébào (shèhuì kēxuébǎn)] Journal of Qingdao Ocean University (Social Science Edition)*. 4, 60-63.
- Lü, C. (2010) The effects of word-knowledge depth, part of speech, and proficiency level on word association among learners of Chinese as a second language. In: Everson, M. & Shen, H. (eds.) *Research among learners of Chinese as a foreign language*. Honolulu, National Foreign Language Resource Center, pp. 67-92.
- Meara, P. (1990) A note on passive vocabulary. *Second Language Research*. 6 (2), 150–154.
- Meara, P. (2004) Modelling vocabulary loss. *Applied Linguistics*. 25 (2), 137-155.
- Meara, P. & Miralpeix, I. (2016) *Tools for researching vocabulary*. Bristol, Blue Ridge Summit, Multilingual Matters.
- Milton, J. (2009) *Measuring second language vocabulary acquisition*. Second language acquisition. Bristol, Buffalo & Toronto: Multilingual Matters.
- Moser, D. (1991) Why is Chinese so damn hard. In: Mair, V.H. (ed.) *Schriftfestschrift: Essays on writing and language in honor of John De Francis on his eightieth birthday*. Philadelphia: University of Pennsylvania Department of Oriental Studies. [Online]. Available from: <http://pinyin.info/readings/texts/moser.html> [Accessed: 8. July 2020]
- Nation, P. (2006) How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*. 63 (1), 59-81.

- Nation, P. (2013) *Learning vocabulary in another language*. The Cambridge applied linguistics series. 2nd edition. Cambridge [u.a.], Cambridge University Press.
- Nation, P. (2012) *The Vocabulary Size Test*. [Online]. Available from: <https://www.wgtn.ac.nz/lals/about/staff/Publications/paul-nation/Vocabulary-Size-Test-information-and-specifications.pdf> [Accessed: 3 March 2020]
- Nation, P. & Stuart Webb (2011) *Researching and analysing vocabulary*. Boston, Cengage Learning.
- Packard, J.L. (2000) *Morphology of Chinese: a linguistic and cognitive approach*. Cambridge [u.a.], Cambridge University Press.
- Perfetti, C.A., Zhang, S. & Berent, I. (1992) Reading in English and Chinese: evidence for a 'universal' phonological principle. In: Frost, R. (ed.) *Orthography, phonology, morphology, and meaning*. Amsterdam [u.a.], North-Holland, pp. 227-248.
- Plag, I. (2003) *Word-formation in English*. Cambridge textbooks in linguistics, 46. Cambridge [u.a.], Cambridge University Press.
- Qi Xuedan [齐雪丹] (2019) 汉语‘词类’之辩 [Hànyǔ 'cílèi' zhī biàn] The Distinction of Chinese word classes. 汉字文化 [Hànzì wénhuà] *Culture of Chinese characters*. 9, 75-82.
- Read, J. (2000) *Assessing vocabulary*. Cambridge [u.a.], Cambridge University Press.
- Read, J. (1995) Refining the word associates format as a measure of depth of vocabulary knowledge. *New Zealand Studies in Applied Linguistics*. 1, 1-17.
- Read, J. (1993) The development of a new measure of L2 vocabulary knowledge. *Language Testing*. 10 (3), 355-371.
- Richards, J. (1976) The role of vocabulary teaching. *Tesol Quarterly*. 10 (1), 77-89.
- Robins, R.H. (1989) Collocation. In: *General linguistics. An introductory survey*. 4th edition. London & New York, Longman, pp. 64-67.
- San, D. (1998) Wordhood in Chinese. In: *New approaches to Chinese word formation. Morphology, phonology and the lexicon in modern and ancient Chinese*. Berlin, New York, Mouton de Gruyter, pp. 134-196.
- Schmitt, N. (2008) Review article: instructed second language vocabulary learning. *Language Teaching Research*. 12 (3), 329-363.
- Norbert Schmitt & Michael McCarthy (eds.) (1997) *Vocabulary: description, acquisition and pedagogy*. Cambridge [u.a.], Cambridge University Press.
- Shen, H. (2008) An analysis of word decision strategies among learners of Chinese. *Foreign Language Annals*. 41 (3), 501-524.

- Shen, H. (2005) An investigation of Chinese character learning strategies among non-native learners of Chinese. *System*. 33, 49-68.
- Shen, H. (2000) Radical knowledge and character learning among learners of Chinese as a foreign language. *Linguistic Studies*. 85-93.
- Shen, H. & Ke, C. (2007) Radical awareness and word acquisition among nonnative learners of Chinese. *The Modern Language Journal*. 91 (1), 97-111.
- Sinclair, J. (1991) Collocation. In: Sinclair, J. (ed.) *Corpus, concordance, collocation*. Oxford, Oxford University Press, pp. 109–122.
- Tian Hailin [田海林] & Xia Jinrong [夏津蓉] (2020) 也谈词的兼类问题 [Yě tán cí de jiānlèi wèntí] A discussion on the categorisation of words. *中国多媒体与网络教学学报 [Zhōngguó duōméitǐ yǔ wǎngluò jiàoxué xuébào]* China Journal of Multimedia and Network Teaching. 4, 239-240.
- Webb, S. (2016) Learning vocabulary through meaning-focused input: replication of Elley (1989) and Liu & Nation (1985). *Language Teaching*. 49 (1), 129-140.
- Weigelin-Schwiedrzik, S. (2019) Email sent to Stephanie Dalton, 19th December. *Betreuung MA-Arbeit*.
- Wesche, M. & Paribakht, T. (1996) Assessing second language vocabulary knowledge: depth vs. breadth. *Canadian Modern Language Review*. 53 (1), 13-40.
- Williams, C. (2013) Emerging development of semantic and phonological routes to character decoding in Chinese as a foreign language learners. *Reading and Writing*. 26, 293-315.
- Woo, J.Y.T. (2004a) Grammatical devices. In: Woo, J.Y.T (ed.) *A linguistic introduction to Chinese languages*. Hong Kong, QESOSA Education Publication. pp. 191-231.
- Woo, J.Y.T. (2004b) Morphemes and compounds. In: Woo, J.Y.T (ed.) *A linguistic introduction to Chinese languages*. Hong Kong, QESOSA Education Publication. pp. 99-126.
- Woo, J.Y.T. (2004c) Word classes. In: Woo, J.Y.T (ed.) *A linguistic introduction to Chinese languages*. 1st edition. Hong Kong, QESOSA Education Publication. pp. 151-190.
- Zhang, D. & Yang, X. (2016) Chinese L2 learners' depth of vocabulary knowledge and its role in reading comprehension. *Foreign Language Annals*. 49 (4), 699-715.
- Zhen Tianyuan [甄天元], Ren Qiulan [任秋兰] & Yin Hailiang [尹海良] (2006) 词语搭配的界定与研究概况 [Cíyǔ dāpèi de jièdìng yǔ yánjiū gàikuàng] Definition and overview of word collocations. *莱阳农学院学报(社会科学版) [Láiyáng nóngxuéyuàn xuébào (shèhuìkēxué bǎn)]* Journal of Laiyang Agricultural College (Social Science Edition). 18 (1), 81-84.
- Zhu Dexi [朱德熙], Lu Jiawen [卢甲文] & Ma Zhen [马真] (1961) 关于形容词名物化问题 [Guānyú xíngróngcí míngwùhuà wèntí] On the problem of nominalisation of adjectives. *北京大学学报 [Běijīng dàxué xuébào]* Journal of Peking University. (4), 51-64.

Appendix A Chinese Word Associates Test (WAT)

*This test is part of my master thesis, which looks at the vocabulary of learners of Chinese as a foreign language. It assesses your ability to identify collocational associates in Chinese. You are going to be asked to decide if any of the four given words form a **collocational association** with the given target word (i.e. if they frequently appear together).*

*Please mark your answer by marking either the **Y** (Yes) or the **N** (No) box next to the word. If you don't know a certain word, mark the box **U** (Unknown). If you do not know a certain target word, you can also mark the box **U** (Unknown) next to it and skip said target word. A typical test item will look like the one below.*

干净（的）				☐ U
意义	☐ Y	☐ N	☐ U	
食物	☐ Y	☐ N	☐ U	
成果	☐ Y	☐ N	☐ U	
房间	☐ Y	☐ N	☐ U	

干净（的）				☐ U
意义	☐ Y	☒ N	☐ U	
食物	☒ Y	☐ N	☐ U	
成果	☐ Y	☒ N	☐ U	
房间	☒ Y	☐ N	☐ U	

In the above example with the target word 干净 (clean) choices 2 and 4 食物 (food) and 房间 (room) would be marked Y (Yes) because they form expressions with the target word, such as “clean food” and “clean room”. Choices 1 and 3 意义 (meaning) and 成果 (achievement) on the other hand would be marked as N (No), as they have no collocational association with the target word.

*You will be given a **maximum time of 25 minutes** to complete the booklet, which contains **45 target words**. After you have completed the booklet, please fill out the demographic questionnaire. Please ask for clarifications if you are unsure about the directions.*

Are you ready to start?

Please turn over the page and

work on the questions in this booklet.

1. 稳定 (的) ☐U

基础	☐ Y	☐ N	☐ U
社会	☐ Y	☐ N	☐ U
答复	☐ Y	☐ N	☐ U
状态	☐ Y	☐ N	☐ U

2. 广泛 (的) ☐U

应用	☐ Y	☐ N	☐ U
关注	☐ Y	☐ N	☐ U
兴趣	☐ Y	☐ N	☐ U
道路	☐ Y	☐ N	☐ U

3. 安全 (的) ☐U

场所	☐ Y	☐ N	☐ U
思维	☐ Y	☐ N	☐ U
成功	☐ Y	☐ N	☐ U
环境	☐ Y	☐ N	☐ U

4. 突出 (的) ☐U

成就	☐ Y	☐ N	☐ U
地位	☐ Y	☐ N	☐ U
贡献	☐ Y	☐ N	☐ U
打击	☐ Y	☐ N	☐ U

5. 特别 (的) ☐U

目的	☐ Y	☐ N	☐ U
含义	☐ Y	☐ N	☐ U
规定	☐ Y	☐ N	☐ U
关系	☐ Y	☐ N	☐ U

6. 明显 (的) ☐U

效果	☐ Y	☐ N	☐ U
空间	☐ Y	☐ N	☐ U
变化	☐ Y	☐ N	☐ U
差别	☐ Y	☐ N	☐ U

7. 高兴 (的) ☐ U

神情	☐ Y	☐ N	☐ U
印象	☐ Y	☐ N	☐ U
样子	☐ Y	☐ N	☐ U
思想	☐ Y	☐ N	☐ U

8. 激烈 (的) ☐ U

竞争	☐ Y	☐ N	☐ U
愿望	☐ Y	☐ N	☐ U
疼痛	☐ Y	☐ N	☐ U
争吵	☐ Y	☐ N	☐ U

9. 方便 (的) ☐ U

成果	☐ Y	☐ N	☐ U
条件	☐ Y	☐ N	☐ U
负担	☐ Y	☐ N	☐ U
功能	☐ Y	☐ N	☐ U

10. 紧急 (的) ☐ U

情绪	☐ Y	☐ N	☐ U
神经	☐ Y	☐ N	☐ U
关头	☐ Y	☐ N	☐ U
任务	☐ Y	☐ N	☐ U

11. 普通 (的) ☐ U

生活	☐ Y	☐ N	☐ U
朋友	☐ Y	☐ N	☐ U
特点	☐ Y	☐ N	☐ U
道理	☐ Y	☐ N	☐ U

12. 简单 (的) ☐ U

单纯	☐ Y	☐ N	☐ U
纯洁	☐ Y	☐ N	☐ U
单调	☐ Y	☐ N	☐ U
容易	☐ Y	☐ N	☐ U

13. 清楚 (的) ☐U

认识	☐ Y	☐ N	☐ U
立场	☐ Y	☐ N	☐ U
眼睛	☐ Y	☐ N	☐ U
目标	☐ Y	☐ N	☐ U

14. 珍贵 (的) ☐U

资料	☐ Y	☐ N	☐ U
礼物	☐ Y	☐ N	☐ U
物价	☐ Y	☐ N	☐ U
友谊	☐ Y	☐ N	☐ U

15. 富裕 (的) ☐U

费用	☐ Y	☐ N	☐ U
家境	☐ Y	☐ N	☐ U
国家	☐ Y	☐ N	☐ U
能力	☐ Y	☐ N	☐ U

16. 成熟 (的) ☐U

技术	☐ Y	☐ N	☐ U
水果	☐ Y	☐ N	☐ U
劳动	☐ Y	☐ N	☐ U
经验	☐ Y	☐ N	☐ U

17. 长远 (的) ☐U

计划	☐ Y	☐ N	☐ U
影响	☐ Y	☐ N	☐ U
水平	☐ Y	☐ N	☐ U
范围	☐ Y	☐ N	☐ U

18. 灵活 (的) ☐U

手指	☐ Y	☐ N	☐ U
反应	☐ Y	☐ N	☐ U
形式	☐ Y	☐ N	☐ U
信息	☐ Y	☐ N	☐ U

19. 痛苦 (的) ☐U

意见	☐ Y	☐ N	☐ U
经历	☐ Y	☐ N	☐ U
选择	☐ Y	☐ N	☐ U
速度	☐ Y	☐ N	☐ U

20. 真诚 (的) ☐U

利益	☐ Y	☐ N	☐ U
祝福	☐ Y	☐ N	☐ U
情形	☐ Y	☐ N	☐ U
帮助	☐ Y	☐ N	☐ U

21. 知名 (的) ☐U

人物	☐ Y	☐ N	☐ U
作品	☐ Y	☐ N	☐ U
表达	☐ Y	☐ N	☐ U
结果	☐ Y	☐ N	☐ U

22. 温暖 (的) ☐U

怀抱	☐ Y	☐ N	☐ U
性格	☐ Y	☐ N	☐ U
阳光	☐ Y	☐ N	☐ U
态度	☐ Y	☐ N	☐ U

23. 出色 (的) ☐U

表现	☐ Y	☐ N	☐ U
人才	☐ Y	☐ N	☐ U
要求	☐ Y	☐ N	☐ U
作品	☐ Y	☐ N	☐ U

24. 细致 (的) ☐U

经历	☐ Y	☐ N	☐ U
服务	☐ Y	☐ N	☐ U
毛病	☐ Y	☐ N	☐ U
分析	☐ Y	☐ N	☐ U

25. 自负 (的) ☐ U

心情	☐ Y	☐ N	☐ U
口吻	☐ Y	☐ N	☐ U
神态	☐ Y	☐ N	☐ U
道德	☐ Y	☐ N	☐ U

26. 锐利 (的) ☐ U

目光	☐ Y	☐ N	☐ U
精神	☐ Y	☐ N	☐ U
时机	☐ Y	☐ N	☐ U
武器	☐ Y	☐ N	☐ U

27. 漂亮 (的) ☐ U

建筑	☐ Y	☐ N	☐ U
竞争	☐ Y	☐ N	☐ U
情感	☐ Y	☐ N	☐ U
衣服	☐ Y	☐ N	☐ U

28. 恰当 (的) ☐ U

事实	☐ Y	☐ N	☐ U
选择	☐ Y	☐ N	☐ U
判断	☐ Y	☐ N	☐ U
评价	☐ Y	☐ N	☐ U

29. 牢固 (的) ☐ U

计划	☐ Y	☐ N	☐ U
身体	☐ Y	☐ N	☐ U
联系	☐ Y	☐ N	☐ U
观念	☐ Y	☐ N	☐ U

30. 热忱 (的) ☐ U

服务	☐ Y	☐ N	☐ U
欢迎	☐ Y	☐ N	☐ U
批评	☐ Y	☐ N	☐ U
希望	☐ Y	☐ N	☐ U

31. 舒畅 (的) ☐ U

线条	☐ Y	☐ N	☐ U
感觉	☐ Y	☐ N	☐ U
气氛	☐ Y	☐ N	☐ U
语言	☐ Y	☐ N	☐ U

32. 详尽 (的) ☐ U

资料	☐ Y	☐ N	☐ U
概括	☐ Y	☐ N	☐ U
介绍	☐ Y	☐ N	☐ U
记录	☐ Y	☐ N	☐ U

33. 充沛 (的) ☐ U

精力	☐ Y	☐ N	☐ U
进步	☐ Y	☐ N	☐ U
资源	☐ Y	☐ N	☐ U
赞扬	☐ Y	☐ N	☐ U

34. 低廉 (的) ☐ U

支持	☐ Y	☐ N	☐ U
成本	☐ Y	☐ N	☐ U
费用	☐ Y	☐ N	☐ U
收入	☐ Y	☐ N	☐ U

35. 公道 (的) ☐ U

说明	☐ Y	☐ N	☐ U
价格	☐ Y	☐ N	☐ U
文易	☐ Y	☐ N	☐ U
成绩	☐ Y	☐ N	☐ U

36. 荒唐 (的) ☐ U

观点	☐ Y	☐ N	☐ U
行为	☐ Y	☐ N	☐ U
成就	☐ Y	☐ N	☐ U
解释	☐ Y	☐ N	☐ U

37. 流畅 (的) ☐U

文字	☐ Y	☐ N	☐ U
讨论	☐ Y	☐ N	☐ U
动作	☐ Y	☐ N	☐ U
声音	☐ Y	☐ N	☐ U

39. 疲惫 (的) ☐U

脚步	☐ Y	☐ N	☐ U
身体	☐ Y	☐ N	☐ U
品德	☐ Y	☐ N	☐ U
道路	☐ Y	☐ N	☐ U

41. 悲伤 (的) ☐U

实际	☐ Y	☐ N	☐ U
失败	☐ Y	☐ N	☐ U
时刻	☐ Y	☐ N	☐ U
情绪	☐ Y	☐ N	☐ U

38. 尴尬 (的) ☐U

局面	☐ Y	☐ N	☐ U
道理	☐ Y	☐ N	☐ U
势力	☐ Y	☐ N	☐ U
滋味	☐ Y	☐ N	☐ U

40. 肤浅 (的) ☐U

眼光	☐ Y	☐ N	☐ U
条件	☐ Y	☐ N	☐ U
体会	☐ Y	☐ N	☐ U
真理	☐ Y	☐ N	☐ U

42. 孤寂 (的) ☐U

世界	☐ Y	☐ N	☐ U
心灵	☐ Y	☐ N	☐ U
事实	☐ Y	☐ N	☐ U
概念	☐ Y	☐ N	☐ U

43. 尖刻 (的) ☐Y☐N☐U

批评 ☐Y☐N☐U

思念 ☐Y☐N☐U

语气 ☐Y☐N☐U

记忆 ☐Y☐N☐U

44. 凌乱 (的) ☐Y☐N☐U

头发 ☐Y☐N☐U

管理 ☐Y☐N☐U

夜晚 ☐Y☐N☐U

房间 ☐Y☐N☐U

45. 圆滑 (的) ☐Y☐N☐U

歌声 ☐Y☐N☐U

色彩 ☐Y☐N☐U

胜利 ☐Y☐N☐U

手段 ☐Y☐N☐U

Appendix B Demographic background questionnaire

This questionnaire gathers demographic background information in order to better classify the results of the Word Associates Test. Please answer the questions below by marking the corresponding box (e.g. ☒ or ☒) or filling in the blanks. When you are unsure of an answer, please give an approximation to the best of your knowledge. Your response to each question is very important for this study and will be greatly appreciated.

Q1. Please indicate your age: _____ years.

Q2. What is your gender?

☐ Female ☐ Male ☐ Other: _____

Q3. Which language is your first language or mother tongue?

☐ German
☐ Chinese
☐
☐ Other: _____

Q4. Do any of your family members whom you grew up with speak Chinese as his or her native language? (If No, skip Q5 and go to Q6.)

☐ Yes ☐ No

Please turn over the page and continue on the back.

Q5. To what extent do you use Chinese to communicate with your family members?

- ☐ Never
- ☐ Occasionally
- ☐ Sometimes
- ☐ Most of the time
- ☐ Almost always

Q6. For how many years have you studied Chinese formally as a foreign language in the classroom?

- ☐ Less than 2 years
- ☐ 2-3 years
- ☐ 3-4 years
- ☐ 4-5 years
- ☐ More than 5 years

Q7. [*Only for MA students*] Did you complete your BA degree in Chinese studies at the University of Vienna? If not, where did you study before you came to study in Vienna?

- ☐ Yes
- ☐ No, I completed my BA degree at _____

Q8. Have you had any experience learning Chinese regularly outside class from a Chinese native speaker? If yes, please briefly describe your experience.

- ☐ Yes, for ____ years ____ months. I learned/used Chinese in the following contexts/activities:

- ☐ No

Q9. Have you ever lived/studied/worked abroad where Chinese is spoken as a native language? If yes, please briefly describe your in-country language related experience.

- ☐ Yes, for ____ years ____ months. I learned/used Chinese in the following contexts/activities:

- ☐ No

Q10. Have you ever taken a standardised test in Chinese? If so, please provide your score and/or the highest level you passed and when you took the test (month/year). If you cannot remember your exact score, please give your best estimate.

- ☐ HSK (Chinese Proficiency Test) _____
- ☐ TOP/TOCFL (Test of Chinese as a Foreign Language)

- ☐ Other:

Q11. Which other languages (apart from Chinese) have you learned? (e.g. in school, in language courses)

- | | |
|----------------------------------|---------------------------------------|
| ☐ English | ☐ Japanese |
| ☐ French | ☐ Korean |
| ☐ Spanish | ☐ Other: _____ |
| ☐ Italian | _____ |

**Thank you very much
for participating in this study!**

Appendix C List of tables, illustrations and figures

List of Tables

Table 1. Types of selection errors (Aitchison, 2012:22).....	17
Table 2. What is involved in knowing a word? (Nation, 2013:49)	47
Table 3. Aspects of vocabulary knowledge for testing (Nation, 2013:538)	62
Table 4. WAT target words and HSK levels	84
Table 5. Language courses and credit distribution of the BA programme	85
Table 6. Definition of solution sheets	92
Table 7. General results WAT (s.sh. 1).....	93
Table 8. General results WAT (s.sh. 2).....	94
Table 9. General results WAT (s.sh. 3).....	95
Table 10. Unknown target words & word options, blank items & word options.....	97
Table 11. Target words marked as unknown by the participants	98
Table 12. Results WAT across frequency levels (s.sh. 1).....	100
Table 13. Results WAT across all frequency levels (s.sh. 2)	101
Table 14. Results WAT across all frequency levels (s.sh. 3).....	103
Table 15. Results WAT: first vs. third semester MA-students (s.sh. 1).....	105
Table 16. Results WAT: first vs. third semester MA-students (s.sh. 2).....	106
Table 17. Results WAT: first vs. third semester MA-students (s.sh. 3).....	107
Table 18. Results WAT: heritage vs. non-heritage learners (s.sh. 1)	109
Table 19. Results WAT: heritage vs. non-heritage learners (s.sh. 2)	110
Table 20. Results WAT: heritage vs. non-heritage learners (s.sh. 3)	111
Table 21. Average and median accuracy rate of the participants' total score	113
Table 22. Accuracy rate by frequency of target word.....	113
Table 23. Accuracy rate of heritage and non-heritage learners	114
Table 24. Results of non-heritage learners (s.sh.1)	116

List of Illustrations

Illustration 1. Example of a Yes/No test item (Beglar & Nation, 2014:173).....	66
Illustration 2. Example of a Vocabulary Levels Test item (Beglar & Nation, 2014:174)	67
Illustration 3. Examples of Vocabulary Size Test items (Nation, 2013: 524).....	68
Illustration 4. Vocabulary Knowledge Scale (Beglar & Nation, 2014:179)	70
Illustration 5. Example of a WAT item (Milton, 2009:162)	72
Illustration 6. Example of a WAT item (Beglar & Nation, 2014:178)	72
Illustration 7. Example of a V_Links measure (Schur, 2007:190)	73
Illustration 8. Example of a V_Links measure (Wolter, 2009:144)	74
Illustration 9. Research questions	75
Illustration 10. Typical WAT item (Chen, 2016:136).....	81

List of Figures

Figure 1. Time spent learning CFL.....	88
Figure 2. Time spent in a country with Chinese as its native language	89
Figure 3. Languages learned by the participants	90
Figure 4. General results WAT (s.sh. 1).....	92
Figure 5. General results WAT (s.sh. 2).....	93
Figure 6. General results WAT (s.sh. 3).....	94

Figure 7. Scatter plots of total scores (s. sh. 1).....	94
Figure 8. Scatter plots of total scores (s.sh. 2).....	95
Figure 9. Scatter plots of total scores (s.sh. 3).....	95
Figure 10. Unknown target words and word options, blank items and word options	96
Figure 11. Results WAT across all frequency levels (s.sh. 1).....	99
Figure 12. Results WAT across all frequency levels (s. sh. 2).....	101
Figure 13. Results WAT across all frequency levels (s.sh. 3).....	102
Figure 14. Results WAT: first vs. third semester MA-students (s.sh. 1).....	104
Figure 15. Results WAT: first vs. third semester MA-students (s.sh. 2).....	105
Figure 16. Results WAT: first vs. third semester MA-students (s.sh. 3).....	107
Figure 17. Results WAT: heritage vs. non-heritage learners (s.sh 1)	108
Figure 18. Results WAT: heritage vs. non-heritage learners (s.sh. 2)	109
Figure 19. Results WAT: heritage vs. non-heritage learners (s.sh. 3)	111

Appendix D Abstract (German)

Diese Studie untersucht die Tiefe des Vokabelwissens von Masterstudierenden der Sinologie an der Universität Wien. Der Fokus liegt hierbei auf der Identifizierung einer bestimmten Art von Wortassoziationen, nämlich Kollokationen, durch die Teilnehmenden. Weiters wird untersucht, wie sich die Häufigkeit der Zielwörter, sowie andere Faktoren (z.B. Sprachlernhintergrund oder Auslandsaufenthalt in China) auf die Fähigkeit der Teilnehmenden auswirken, Kollokationen zu identifizieren. Darüberhinaus vergleicht diese Studie auch verschiedene Gruppen von Lernenden unter den Teilnehmenden (d.h. Masterstudierende des ersten und dritten Semesters, sowie Lernende mit chinesisch-sprachigem Hintergrund und Lernende ohne solch einen Hintergrund).

Zu diesem Zweck wurde ein Word Associates Test durchgeführt. Dieser wurde durch einen Fragebogen zum demografischen Hintergrund der Teilnehmenden ergänzt, um die Ergebnisse besser einordnen zu können. Es wurden Daten von sechzehn Teilnehmenden gesammelt. Diese nahmen zum Zeitpunkt der Datenerhebung an einem Sprachkurs des Masterstudienprogramms teil.

Die Ergebnisse der Studie zeigen, dass die Teilnehmenden eine hohe Anzahl an Zielwörtern kennen, aber ihre Fähigkeit Kollokationen zu identifizieren hinkt hinterher. Außerdem scheinen andere Faktoren, die oben erwähnt wurden, keinen sichtbaren Einfluss auf die Ergebnisse der Teilnehmenden zu haben.

Die geringe Teilnehmerzahl stellt eine Einschränkung dieser Studie dar, da die Ergebnisse möglicherweise nicht repräsentativ sind. Trotzdem lassen sich aus diesen Ergebnissen nützliche Implikationen über das Wissen von Lernenden von Chinesisch als Fremdsprache über chinesische Kollokationen ziehen.