A quantitative study of Swedish high school L2 learners' receptive knowledge of English lexical collocations

Louise Frifelt
ENGK06
HT23
Supervisor: Henrik Gyllstad
Centre of Languages and Literature

Abstract

This paper investigates receptive knowledge of English lexical collocations among Swedish high school students. The study tests the claim that L2 (second language) learners of English have a low level of collocational competence and tend to recognise high-frequent collocations more than low-frequent since the former have been encountered more. The target items for this study consisted of 40 collocations with various frequencies and combinations of Verb+Noun (V+N) and Adjective+Noun (V+N) since previous research indicates that the latter is even more difficult for L2 learners.

The participants were 56 high school students, L2 learners of English, between the ages of 16-19 from Malmö and Lund in southern Sweden. They were instructed to fill out a questionnaire in a form recognition format, i.e. multiple-choice, where they were tasked to pick the most natural combination of words in English out of three alternatives.

The results showed an advanced level of receptive collocation knowledge among Swedish upper secondary pupils, with a mean score of 73,1%. A possible explanation could be that Swedish pupils are highly proficient in English and spend a considerable amount of time on Extramural English (EE), since vocabulary is given lower priority in the syllabus.

In addition, the pupils also displayed higher knowledge of V+N than Adj+N collocations.

On top of that, the students also knew high frequent collocations to a greater extent than low-frequent collocations, generally and within both structures (V+N and Adj+N).

Another purpose of this essay was to examine the relation of speaking, listening, reading and writing with collocation knowledge. The findings suggested a possible connection between collocational competence and reading and writing. Further research within this aspect is recommended.

Acknowledgements

First off, I would like to thank Satu Manninen, the course moderator for this linguistic BA project course. Your words during a lecture would later give me inspiration for my research question in this thesis.

Special thanks to my supervisor Henrik Gyllstad, whose expertise and guidance has helped me greatly throughout the writing of this essay.

Thank you to Ann Kull at the academic support centre for giving me general advice, tips and tricks and sometimes being there to brainstorm ideas.

I would like to thank my fellow bachelor classmate Maria for helping me gather participants, and high school students Sif and Malwa, for participating in my survey as well as helping me gather more participants. Without you, I would not have any data to collect. You deserve a big thank you!

A big thank you to my friends, Victoria, Martin, Kenny, Dennis, Natalie, David, Jinhee and Jennifer. Not only did you support and encourage me a lot from start to finish, but you were also my study buddies and offered input and advice for this type of study.

Last but not least, I want to thank my mother Renée and my father Christian, for encouraging and pushing me towards and through my education. Without you I would not be doing this today. Thank you!

Table of contents

1. Introduction

When it comes to second language learning, vocabulary was sidelined from the curriculum since there is little focus in literature about second language acquisition in general. Milton (2009) argues that this originates from structural linguistics, as well as other approaches to language teaching that have become more prevalent. Outside specialist studies in vocabulary, as an academic or pedagogical issue, vocabulary is unchallenging; there seems to be the perception that learning words is unsystematic because words are just words. Moreover, in structural methods towards learning, the important part is how systems and rules are acquired and as a consequence there is no need to be concerned about the words these rules apply to (p. 1).

According to Meunier and Granger (2009), research in psycholinguistics has shown that language is learned, processed as chunks or blocks. As a consequence, it would be illogical to not suggest that phraseology should fill a central and unquestionable position in "instructed second language acquisition" (p. 247). Moreover, one type of multi-word unit (MWUs) is collocation, "a natural combination of two or more words", as in *do homework* instead of **make homework* (McCarthy & O'Dell, 2017, p. 4). Barfield and Gyllstad (2009) point out that the importance and concern of teaching collocations have been observed by researchers for the past decades. The interest in researching the function and nature of collocation has risen, given the recent increased interest in vocabulary studies. Research based on learner corpus has shown that although L2 (second language) learners use and recognise fewer collocations, they also utilise fewer collocations compared to L1 (first language) speakers. Furthermore, classroom-based research has suggested that the knowledge of collocations' meaning and form among L2 learners is less than that of single-word items (Newton et al., 2013).

Collocations are an essential part of the everyday language since they are one of the keys to mastering fluency and a higher level of English (McCarthy & O'dell, 2017, p. 4). However, a considerable amount of studies show that not only is collocation knowledge amongst L2 learners of English low, even so at an advanced level; but the development and learning process is also slow. Moreover, there are several distinct combinations of collocations, but Verb+Noun and Adj+Noun in particular are shown to be more difficult for nonnative speakers of English (Nguyen & Webb, 2017).

Information on prefabricated units and multi-word combinations is limited. Research on Swedish L2 learners of English in particular is even more scarce. However, proficiency in

English is said to be at a high level among Swedish learners and Swedish pupils in particular. Thus, I considered the main research question:

- How is Swedish high school students' receptive knowledge of lexical collocations?

In addition to the main research question, I choose to also examine the connection between collocation knowledge and reading, writing, speaking and listening.

The target collocations included in the questionnaire will be based on frequency of co-occurrence. The test items are Verb+Noun (V+N) and Adjective+Noun (Adj+N). There has not been much research done on Swedish L2 learners' receptive knowledge of collocations, let alone Swedish high school students. Therefore, this study will fill in a gap in knowledge on this topic. Based on previous research (Nguyen & Webb, 2017), the hypothesis is that the students will know the collocations of the highest level(s) of frequency and the V+N collocations better than the Adj+N collocations. Additionally, this study might also confirm another finding that L2 learners of English have a low level of collocation knowledge.

2. Background

Section 2 Background will explain theoretical approaches, clarify the criteria for collocation in this study and present previous research concerning L2 learners' collocational competence, L2 acquisition and L1 congruence. Swedish pupils' skills and attitudes towards English will also be mentioned.

2.1. Phraseology

Language is naturally formulaic and contains fixed expressions, also known as multi-word items (MWIs) (Snoder, 2019, p. 1). These are prominent in most of the world's languages besides "single orthographic words", which are letters ringed by space (Gyllstad, 2020, p. 387). There are several common multi-word structures in English, namely compound nouns, phrasal verbs, idioms and collocations. The last-mentioned structure will be the focus of this essay. The definition of collocation is far from well-defined, but several attitudes and methods have been used in previous research (Gyllstad, 2007, p. 6). There are two main approaches for defining collocations that are ubiquitous in literature, namely phraseology and frequency.

Not only is phraseology, the study of word combinations, a fairly new field; it is currently gaining attention since a remarkable growth in interest over the last recent decades

can be noticed. It is an interdisciplinary perspective, including several different areas, as opposed to solely linguistics (Meunier & Granger, 2008). In this viewpoint, collocations are characterised by restricted co-occurrence and substitutability; a combination like *good fortune* is regarded as a collocation, but *fortune* does not collocate with *nice, wonderful* or *pleasant* (Nguyen & Webb, 2017, p 299). In other words, collocation is a particular type of "word combination, displaying various degrees of fixedness" (Gyllstad, 2007, p. 15). A division is made between free combinations and collocations. In the first unit, the semantic properties regulate the degree of distinction. Contrastingly, in collocations, the degree of distinction is somewhat arbitrary. *Read a newspaper* cannot be substituted with *drink a newspaper* or *read water* since *read* must take a noun with the semantic property of *containing written language*. Likewise, *drink* requires a noun denoting *liquid*. Conversely, in *reach a decision*, the noun can be replaced with *goal, compromise, verdict* and *conclusion* and this restriction is a result of the language's arbitrary convention rather than semantic properties (Nesselhauf, 2003, p. 225). As a result, there are different types of multi-word units. Nesselhauf (2003) defines three distinct multi-word combinations as follows:

> **Free combinations** (e.g. *want a car*):
> The senses in which the verb and the noun are used are both unrestricted so they can be freely combined.
> **Collocations** (e.g. *take a picture*):
> The sense in which the noun is used is unrestricted, but the sense of the verb is restricted, so that the verb in which it is used can only be combined with certain nouns (*take a picture/ photograph* but e.g. **take a film/movie*).
> **Idioms** (e.g. *sweeten the pill*):
> Both the verb and the noun are used in a restricted sense, so substitution is either not possible at all or only possible to an extremely limited degree. (Nesselhauf, 2003, p. 226)

Consequently, the verb is the distinguishing component between collocations and combinations and likewise, the noun distinguishes collocations from idioms (Nesselhauf, 2003). These specific types of word combinations can be further explained by Howarth's continuum model, a scale of "lexical composites" (Howarth, 1998, p. 28). Firstly in free combinations, e.g. *blow a trumpet*, the components are all used in their literal sense and can be substituted without affecting the meaning. Secondly, in restricted collocations, such as *blow a fuse,* one component word is used in a figurative or delexical sense only with a limited set of collocates, while the other is used in a literal sense. Thirdly, figurative idioms, e.g. *blow*

your own trumpet, have both a literal interpretation and metaphorical meaning. Lastly, in pure idioms, such as *blow the gaff*, the meaning cannot be determined from each constituent since the combination as a whole has a complete meaning, making it the most fixed and semantically opaque classification (Howarth, 1998).

2.2. Frequency

Another prominent approach to study collocations is frequency, even called the statistical (Greenbaum, 1974, p. 83) and textual standpoint (Barfield, 2013, p. 3). Nguyen & Webb (2017) define collocations as "word combinations that co-occur more frequently than could be expected by chance" (p. 299) or "co-occurrence of words in a certain span" (Nesselhauf, 2003, p. 224). Subsequently, in order to demonstrate which multi-word combinations occur with higher and lower frequency, studies in this perspective are likely to include investigation and statistical data-analysis in language corpora (corpus in singular) (Nguyen & Webb, 2017, p. 299), which are "large collections or databases of language, incorporating stretches of discourse ranging from a few words to entire books" (Schmitt, 2000, p. 68). They contain real and authentic data that allows learners to use natural language instead of made up examples and intuitions (Schmitt, 2000). Researchers can also use other statistical measures such as mutual information scores, t-scores, z-scores or mutual information scores. Not only do these methods demonstrate which combinations have a high frequency, but also quickly identify a word's collocates (Webb et al., 2013, pp. 92-93); a search in Corpus of Contemporary American English (COCA) shows that *take* has a high mutual information score with *agreement, conclusion* and *consensus*, suggesting their co-occurrence far more than could be expected.

2.3. Definition of collocation

As mentioned in section 2.1, collocations are a particular type of "word, combination, displaying various degrees of fixedness" (Gyllstad, 2007, p. 15). But there are numerous definitions for this notion. According to Snoder (2019), collocations are defined as a subset of prefabricated units, i.e. fixed phrases (p. 17). Schmitt (2000), defines the notion of collocation as "the tendency of two or more words to co-occur in discourse" (p. 76). There is a systematicity, which is largely linguistic, affecting the choice vocabulary in discourse; fluent English speakers know the correct way of saying *to set something on fire,* although there is no reason why *to put something on fire* is incorrect. Schmitt mentions another key element,

namely the different levels of exclusivity of the relationships between the two items; an example is the word *blonde* which collocates particularly with *hair*, but also with *lady* or *woman*. Contrastingly, it never co-occurs with *wallpaper* or *paint*, even though there is no semantic reason why it is incorrect, and can thus be said to "collocate strongly". Nevertheless, there are also combinations that "collocate weakly". For example, *nice* appears together with any noun associated with pleasantness, e.g. *nice car*, *nice salary* or *nice view* (p. 77).

The vast majority of authors agree on two types of collocations, namely grammatical and lexical collocations. The first type consists of a dominant word that fits together with a grammatical word, usually a verb, noun or adjective succeeded by a preposition, e.g. *access to*, *acquainted with* and *abide by*. On the other hand, lexical collocations usually contain two practically equal words and the meaning of the combination is derived from both words. Examples of this include verb+noun (*spend money*), noun+verb (*ball bounces*) and adjective+noun (*cheerful expression*) (Schmitt, 2000, p. 77).

Barfield (2013) describes lexical collocations as multi-word units consisting of two or more words that form a specific meaning and are frequently used in speech and writing by discourse communities (which are groups that share common goals and use communication to achieve their goals). Lexical collocations are regarded as an important means to sustain fluency, create meaning and speak the language with an appropriate quality of naturalness since they have a status as conventionalized phrases that are repeated regularly (p. 1).

The criteria for the target collocations in this essay are Adj+N and V+N. In the latter, the combination Verb+determiner+Noun is also included (e.g. *take a picture*). As a consequence, lexical collocations are the main focus of this essay. On top of that, differing frequency of co-occurrence in a language corpus is also a factor in choosing these test items.

2.4. Collocation knowledge among L2 learners of English

As mentioned in the introduction, research on collocation among learners of English is currently in its infancy, especially on Swedish learners. Therefore, studies performed in other countries than Sweden are also included in this section.

Collocation is strongly associated with advanced learners of English; as their proficiency increases, they are more likely to use collocations (Saudin et al. 2017, p. 189). However, many researchers agree that L2 students of English have difficulties with collocations in acquisition as well as usage (Begagić, 2014: Nesselhauf, 2003), especially compared to native speakers (González Fernández & Schmitt, 2015, p. 96). Not only is

knowledge of collocation among L2 learners low; but they also tend to process knowledge at a moderate rate. Nguyen & Webb (2017) found that the proficiency was far away from a mastery level since the participants knew less than 50% of the tested collocations. Their knowledge abated at every level of frequency. Learners often use combinations incorrectly or make mistakes. Nesselhauf (2003) found that the most common mistake in a V+N collocation involves a wrong choice of verb, e.g. *carry out races* instead of the correct *hold races*. This may be due to the verb's restricted sense in collocation. Furthermore, the concept of collocational teddy bears involves ESL (English as a second language) and EFL (English as a foreign language) learners, despite their advanced level of English, using an insufficient amount of adjectives. As a result, they produce incorrect and awkward collocations (Nesselhauf, 2005, p. 69).

But some studies illustrate that learners usually are familiar with more collocations than generally thought. González Fernández and Schmitt (2015) investigated productive knowledge of general lexical collocations among Spanish EFL learners. In the results, the participants produced a mean score of 56.6 % of correct answers, indicating that L2 students of English have substantial collocation knowledge.

Studies have also shown that some particular types of collocations are notably more troublesome for L2 learners, particularly Verb + Noun (V+N) and Adjective+Noun (Adj+N) structures. Nguyen & Webb (2017) investigated the knowledge of adjective-noun and verb-noun collocations at levels of the first three 1000 word frequency; although the above-mentioned types of structures were found to be more frequent, they were also precarious for learners. Not only did the results show that the participants had inadequate knowledge of collocation after "formal language instruction" (p. 317) of at least seven years, but also that there was not much difference between knowledge of verb-noun and adjective-noun collocations (Nguyen & Webb, 2017). In many instances of V+N collocations, as in *make progress* and *take care*, meaning is not added to the entire combination by the verb (Gyllstad, 2009; Nguyen & Webb, 2017). In addition, what was notably difficult for learners were causative uses, meaning that the verb causes something to happen, e.g. *make a decision* and *make something possible,* and delexical uses, which is when a verb has little to no meaning by itself, e.g. *make a decision* and *make progress* (Nguyen & Webb, 2017, p. 300).

Previous studies have indicated a strong correlation between collocation knowledge and single-word item vocabulary (Nguyen & Webb, 2017). Masrati (2023) found that collocation skills seemed to develop with increased single-word vocabulary. Another finding was that collocational competence knowledge had a larger significance for L2 proficiency

than single-word items. Gyllstad (2007) conducted a study where he developed tests to measure receptive collocation knowledge. Another purpose was to investigate the relationship between Swedish L2 learners' receptive knowledge of collocations and their vocabulary size. Lastly, the final question was to examine the relationship between the learners' learning level of English and their receptive knowledge of collocations. The results showed that scores from the two tests (COLLEX and COLLMATCH) correlated strongly with single vocabulary size and proficiency level and that the most advanced learners' receptive collocation knowledge matched that of native speakers (Snoder, 2019). The data also showed that "receptive collocation knowledge, as measured in COLLEX and COLLMATCH, essentially increase[d] as a function of higher learning level" (Gyllstad, 2007, p. 251); Swedish University students outperformed Swedish High school students. However, there was a difference between University students with learning levels one year apart (Gyllstad, 2007).

According to Saudin et al. (2017), linguists classify collocational knowledge (or competence) into productive and receptive knowledge and the benefit of this division is that the precise level of collocational competence among learners of English can be measured more accurately. Saudin et al. argue that productive knowledge is more important than receptive knowledge since it is of greater difficulty for EFL learners; the skill requires explicit knowledge that some particular words can collocate together whereas others can not. Moreover, it is related to productive skills such as speaking and writing, whereas receptive collocation knowledge involves recognition of two often co-occurring words in reading and listening (pp. 191-192). They investigated Indonesian EFL learners' and the analysis revealed that their receptive collocational knowledge was higher than their productive knowledge. Moreover, it was also found that the productive and receptive competence scores of V+N collocation were higher than those of Adj+N collocation, suggesting that Adj+N collocations are more difficult (Saudin et al. 2017).

2.4.1. L2 acquisition of collocation

Although it was fairly unclear how collocations are best learned (Schmitt, 2000, p. 79), McCarthy and O'Dell (2017) point out the importance of learning collocations. Not only would they help increase your vocabulary, but they are also a key to speaking and writing a language more naturally and more accurately (p.4). Collocations are therefore regarded as a fundamental element in learning a language; how meaning is created by an amalgamation of words is crucial for all use of language (Barfield & Gyllstad, 2009).

According to Begagić (2014), language theories continue to argue that the main determining factor for acquiring multi-word items is the frequency of occurrence of linguistic elements (in an input); learners of English tend to know high-frequency items more since they are more likely to encounter them. Begagić compared first-year and fourth-year students of English at the University of Zenica in Bosnia and Herzegovina. The results revealed that the latter group had a score higher than the first group, indicating that the senior students had been exposed more to collocations through course literature and lectures. Newton et al. (2013) studied how repetition affects learning of collocations, and found that collocations can be learned incidentally through repetition. Additionally, it appeared that L2 learners do not acquire as many low-frequent collocations (González Fernández and Schmitt, 2015). Previous research illustrates that learning and the processing rate of English multi-word combinations by L2 learners are affected by the frequency level of collocations (Nguyen & Webb, 2017).

According to Snoder (2019), literate L2 learners process multi-word units analytically, i.e. by breaking combinations down into single words, e.g. *make + a + mistake*. Moreover, when the need arises to reassemble the meaningful units of the concept they wish to convey, other semantically motivated candidates may be deemed just as appropriate, such as **do a mistake*. This implies that collocations may not be intrinsically formulaic for L2 learners. Another semantically motivated candidate, e.g. **do an error*, could be considered as appropriate if it was necessary to reassemble the meaningful parts of a concept they wanted to describe. It also means that collocations are not necessarily formulaic for L2 learners (p. 22).

Gyllstad and Wolter (2016) compared the difference in processing collocations and free combinations and found that the former combination was processed slower, most likely due to the semi-transparent quality of collocation. In addition, the results also suggested that phrasal frequency and semantic transparency were key factors in processing collocations.

Exposure to English, inside the classroom as extracurricular activities like social networking sites and movies, is likely to improve the acquisition of L2 skills and collocations (Gyllstad, 2007), especially after more than several years of studying English (González Fernández & Schmitt, 2015). All kinds of contact with a target language outside the classroom is known as Extramural English (EE) (Sundqvist, 2010, p. 95). An explanation for L2 learners' lack of collocational competence can be how they are taught in school; The focus of vocabulary teaching is on single words contrary to multi-word combinations. Students therefore tend to define individual words. They might also be less exposed to multi-word units and may consequently misunderstand the meaning of these words in context (Nguyen & Webb, 2017, p. 313). Also, students are often taught grammar rules instead of vocabulary

(Begagić, 2014, p. 48). Moreover, active vocabulary study is marginalised in Swedish schools, given the action-oriented approach to language (Snoder, 2019, p. 7) and the emphasis on the ability to adapt language to different contexts as well as different ways of communication, i.e. speaking and writing (Skolverket, 2011).

Several sources argue that collocations are worthy of greater attention in teaching than received thus far (Nesselhauf, 2003, pp. 237-238). Due to the recent consensus that knowledge of collocations is a crucial prospect in learning a language (Webb & Kagimoto, 2011, p. 259), both teachers and learners should pay more attention to collocations (Nguyen & Webb, 2017, p. 317). Increased collocational knowledge will not only lead to an enhanced accuracy of a language; but it also facilitates pragmatic skills and fluency (Webb & Kagimoto, 2011, p. 259).

2.4.2. L1 influence and congruence

L1 congruence signifies "word for word overlap between L1 and L2 form and meaning" (Webb & Kagimoto, 2011, p. 265). The reason for English language students' inadequate collocation knowledge may be because collocations in the target language conflict with those in their mother tongue (Begagić, 2014, p. 48). In other words, lack of L1 and L2 congruence may be a significant barrier in processing L2 collocations (Gyllstad & Wolter, 2016, p. 302) and even advanced learners tend to make interlingual errors (Laufer & Waldman, 2011). Nesselhauf, (2003) carried out a study on German-speaking students of English and found that the L1 had a significant influence (45%) on the mistakes, of which collocations were the most affected. This indicates that L1 influence has a strong impact on collocational competence and use.

Another obstacle to learning L2 collocations is the lack of word-for-word translational overlap between the learner's L1 and the L2 in question. For example, the English verb-noun collocation *keep a diary* corresponds to *föra dagbok* in Swedish (literally 'conduct diary'), and not the literal translation from English *hålla en dagbok (Snoder, 2019, p. 24).

Shaw and Wang (2008) compared L1 Swedish and L1 Chinese advanced learners' English collocational errors to test this hypothesis. Not only were the errors similar, despite Swedish and Chinese being extremely different typologically; the participants also produced similar sets of noun collocates. Thus, the authors further concluded that intralingual problems may also be an explanation for collocational errors.

2.5. Swedish pupils' proficiency in English

Studies have shown that Swedish pupils are generally (highly) proficient in English (Sylvén, 2018) and they have strong confidence in their linguistic skills as well. Swedish adolescents have a positive attitude towards the language as a school subject and are exposed to English daily inside and outside of school. The latter activities include travelling, watching TV, movies or videos in English as well as spending time on the internet and playing computer games. In addition, an extensive amount of time is also spent listening to music in English (Erickson, 2004, p. 47). Extramural English (EE) has also been stated to be of significant importance for skills in English (Sylvén, 2018, p. 31). Sundqvist (2010) examined the effects of EE on vocabulary and oral proficiency of Swedish ninth graders. The results showed that productive extramural activities (reading, computer games and surfing the internet) rather than passive activities (watching TV and listening to music) had a stronger impact. In addition, boys had a larger vocabulary than girls; in contrast to girls, boys spent considerably more time on productive extramural activities.

English as the language of instruction in Swedish schools might be another reason for strong English skills among Swedish students; compared to the other countries mentioned in a report by Sylvén (2018) (Finland, Germany, the Netherlands and Spain), where the subject is taught in English at 20-50 %, almost all teaching is done in English. However, Olsson (2016) claims that exposure to English affects colloquial (everyday) language specifically since academic words are more prominent in high school.

3. Methods

3.1. Combining phraseology and frequency

Gyllstad (2007) points out various studies where researchers have differed from either the school of phraseology or the school of frequency since they combined these two aspects (pp. 16-17). The approach for this essay was a combination of phraseology and frequency; the purpose of the study is to investigate Swedish L2 learners' receptive knowledge of collocations. However, the last-mentioned perspective solely underlines any word combinations co-occurring more frequently (Gyllstad, 2007; Nguyen & Webb, 2017, p. 300). For example, *any damage* would be considered a collocation since this particular word combination is found in COCA, but given that the target collocations in this study are Verb+Noun and Adjective+Noun structures, a more fitting word combination in this case would be *cause damage*. On the other hand, the phraseology approach defines different types

of fixed phrases and collocation, which is a type of fixed expression in the perspective of phraseology is specifically explored in this study.

Since the purpose of this study is to find patterns and to investigate the collocation knowledge of a larger population, the data collection method was a quantitative survey. Quantitative research is positive because it focuses on gathering numerical data, prioritises objectivity, and generalisation can be done across groups of people to explain a particular phenomenon (Babbie, 2010, as cited in Research guides: Organizing your social sciences research paper: Quantitative methods, n.d.).

3.2. Designing the questionnaire

For the purpose of anonymity, accessibility, convenience and to facilitate handling the data, the questionnaire was designed to be taken online utilising google forms. Given Swedish pupils' high proficiency in English, I decided to not construct the survey in Swedish. Firstly, the students were asked to fill in their age and gender. Secondly, their native languages exclude native English speakers. Thirdly, the following questions sought to investigate the participants' estimated proficiency in speaking, listening, reading and writing in English. Fourthly, the pupils were requested to indicate the number of hours they used English weekly. These questions were included to further answer my research questions, provide extra aspects and possibly pave the way for future research. Two practice questions, which the participants were asked to answer before being tested on the actual target collocations, were presented to demonstrate the task format. Afterwards, they would be tested on the target collocations of this study.

There are four different formats for testing collocations known as form recognition, form recall, meaning recognition and meaning recall (Gyllstad, 2020). The purpose of this study was to measure receptive L2 collocation knowledge, instead of productive knowledge, without taking L1 form and meaning into account. Hence, I designed the questionnaire in a form recognition format. The students were presented with three alternatives and their task was to tick the answer they believed was correct to test whether they recognised the correct English collocation or not (See Appendix B for more information).

The design of form recall proves more of a challenge; there is a risk that the test taker is unfamiliar with the target multi-word unit and as a consequence more difficult to guess. On top of that, correct answers on this test variety do not necessarily ensure that the participants can produce these kinds of phrases in speaking and writing, i.e. natural language communication (Gyllstad, 2020, p. 391; Cloze test in English: Expert tips to solve accurately

with solved examples. 2021). Furthermore, meaning recognition and meaning recall are not suitable since the relation between L2 form and L1 meaning is not accounted for in this study.

I randomised the order of the different options for each question manually: a randomisation setting in my digital questionnaire was not possible since the consent, questions about age, gender and proficiency of different aspects in English would consequently also be included in the randomisation.

3.3. Choice of target collocations

The target collocations for this study were picked from *English Collocations in Use* (McCarthy & O'dell, 2017) in order to obtain as diverse collocations as possible. They were displayed according to different topics (weather, time, sport etc.). The collocations presented in the book were selected from learner corpora according to usefulness in written and spoken language, not being "immediately obvious" (p. 4) (to put it another way, semi-transparent), and tendency to be problematic for learners, and therefore suitable for this study.

Afterwards, I turned to the *Corpus of Contemporary American English* (COCA) and searched the frequency of each collocation as well as the components of them. I picked out collocations from chapter to chapter until I had 40 of them (20 Adj+N and 20 V+N). However, during the stage of analysing the collected data, I discovered a couple of mistakes concerning the distractors and correct collocations in the questionnaire. Consequently, 36 collocations (18 V+N and Adj+N respectively) were included in the final analysis. They were chosen based on their frequency of co-occurrence in COCA, thus including collocations with various frequencies. In this study, the upper half, with 15938–891 appearances in COCA, was defined as high-frequent. The lower half, with 619–11 appearances in COCA, was defined as low-frequent.

The frequency approach was also used to test what previous research has indicated, namely that L2 learners tend to know high-frequent collocations. To limit the test items further, I have focused on V+N and Adj+N structures; previous research has proven that the foregoing structures tend to be particularly difficult for L2 learners of English (Nguyen & Webb, 2017; Webb et al., 2013). Thus, it would be interesting to test these structures on Swedish L2 learners of English.

3.4. Participants

To get hold of participants for my study, I contacted 20 high schools in Malmö, Lund and Trelleborg out of convenience. Two schools, one in Malmö and one Lund were interested in participating. To protect their integrity, there was no information collected on which school each student was attending. All in all, 56 students answered the questionnaire. Three had to be discarded from the analysis since those participants had English as their native language, and this study focuses on L2 learners of English. Among the 56 students, 34 of them had Swedish as their L1, 5 had Norwegian, 4 had Danish and German, French, Polish and Russian had one native speaker each. In addition, two pupils had multiple native languages. One had Swedish and Danish. The other had Swedish, English and Arabic.

The students were between the ages of 16-19. 22 participants were 16 years old, 17 were 17, 12 were 18 and 5 were 19. However, due to the limited age range as well as the amount of students from each age, this variable was not included in the final analysis. Out of 56 informants in total, 30 were girls, 23 were boys and 3 non-binary, as illustrated in Figure 1. This distribution of girls, boys and non-binary is representative of the population and conclusions can consequently be drawn.
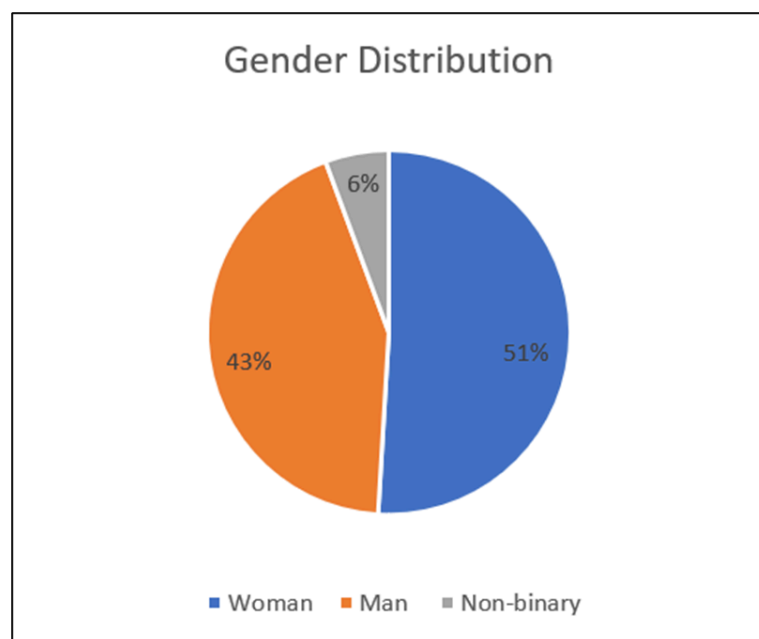

Figure 1 presents the distribution in the percentage of genders that answered the questionnaire.

The choice of high school students in particular was out of interest; collocation is an advanced form of vocabulary knowledge (Schmitt, 2000) and we can thus assume that the higher the grade, the more collocation knowledge one has. This in combination with the

consensus that L2 learners of English have a low level of collocation knowledge (González Fernández & Schmitt, 2015; Nguyen & Webb, 2017) will provide an interesting analysis.

4. Results and discussion

4.1. General knowledge level

Here I will present the findings and descriptive statistics over the collected data based on the answers from the questionnaires. The primary purpose of this study was to investigate the receptive collocation knowledge among Swedish upper secondary school pupils. They were presented with a questionnaire in a multiple-choice format and the task was to choose one answer that they thought was the most natural combination.

The participants produced a mean score of 73.1 %, indicating that they know a considerable amount of English collocations. The scores ranged between 39-94 %. Interestingly, 50 out of 53 students answered correctly on more than half of the target collocations, with most receiving a mean score between 70-80 % as illustrated in Figure 2 below.
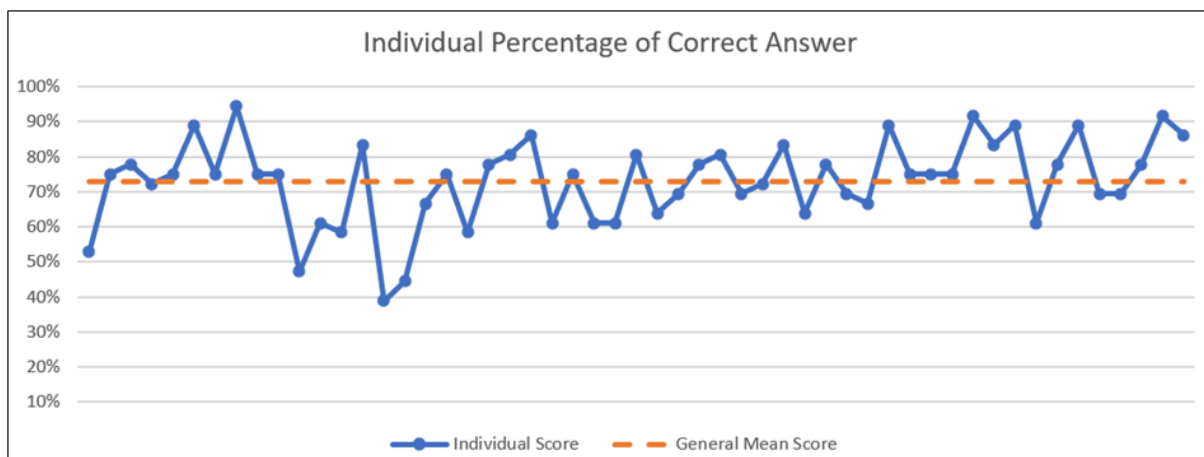
Figure 2 presents each student's total score in percentage together with the general mean score.

These findings contrast with the consensus from previous research, that the collocation knowledge among L2 learners of English is poor (Nguyen & Webb, 2017). Moreover, their high knowledge level is also supported by the study carried out by González Fernández and Schmitt (2015), which demonstrated that a mean score of 56.6 % is an implication of substantial collocation knowledge.

The participants' strong level of collocation knowledge may be due to L1 congruence and influence. Swedish and English are closely related languages. For example, *take a picture* and *play games* are the same in Swedish ('ta en bild', 'spela spel'). By contrast, the least frequent collocations in this study, e.g. *tranquil countryside* and *decent meal*, are not congruent in Swedish ('lugn landsbygd', 'anständig måltid').

Another possible explanation for the students' immense collocation knowledge can be the amount of exposure and usage of English outside the classroom. When asked how many hours they use English per week, five participants even specified that they use English for various activities. Those include reading, computer games, social media, chatting with friends online, watching videos, television shows and even part-time jobs. Furthermore, as mentioned in the background, it has been stated that frequency is the biggest factor for acquisition; one is more likely to know collocations the more one has encountered them earlier (Snoder, 2019). Thus, a possible interpretation of the participants' high scores could be that they have been exposed to (these) collocations before.

As mentioned in the background, vocabulary is generally sidelined in the curriculum (Milton, 2009, p. 1) especially in Sweden (Skolverket, 2011). Thus, we can assume that extracurricular usage of English has a larger impact on Swedish students' collocation knowledge compared to school.

4.2. Which of V+N and Adj+N was best known?

Differences in the participating students' receptive knowledge of V+N and Adj+N collocations were also further investigated and compared in this study. One reason is that the latter structure has been suggested to be considerably difficult for L2 learners of English (Snoder, 2019). Another reason was in consideration of the phraseology perspective, which underlines several different types of set phrases (free combinations, idioms and collocations) and additionally multiple distinct combinations within collocations, with V+N and Adj+N being the main focus of this study.

Figure 3 shows that the scores were generally higher among the V+N collocations than in Adj+N. The mean score among the two combinations differs by 8%; Adj+N was 69%, whereas it was 77% in V+N.
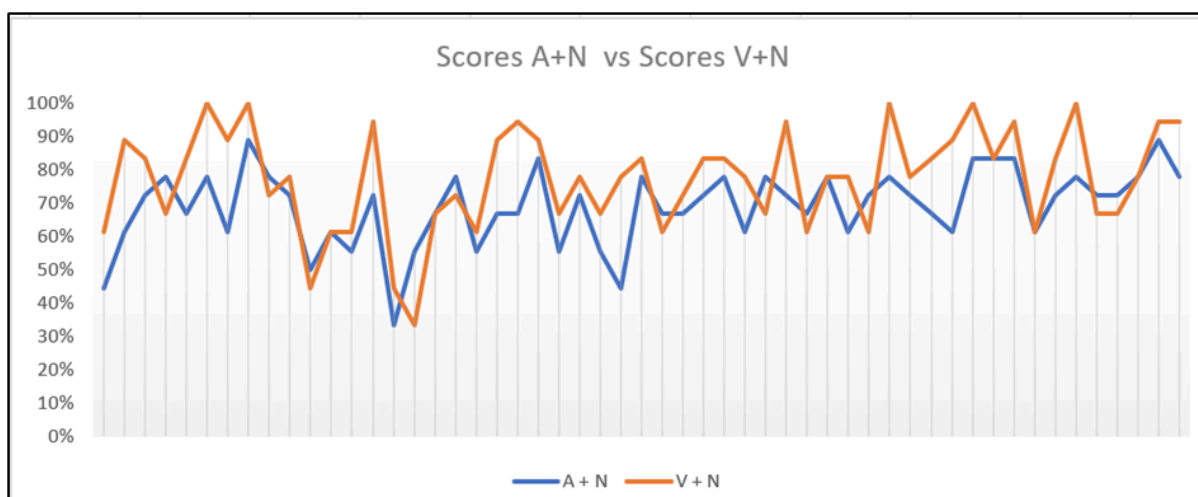
Figure 3 illustrates the total scores in percentage of V+N and Adj+Noun.

An interesting finding is that in five V+N collocations (e.g. *play games*, *single parent*), the mean score was astoundingly 100%. Contrastingly, there was no Adj+Noun combination where every student answered correctly.

In only 6 out of the 36 target collocations, fewer than half of the participants answered correctly. Those were the *tranquil countryside, decent meal, sole survivor, vivid imagination, review books* and *take an exam.* As can be seen, the majority of these six collocations are Adj+N combinations and this strengthens the claim of V+N being less difficult for L2 learners (Begagić, 2014; Saudin et al., 2017).

4.3. High frequency vs low frequency collocations

The aspect of frequency was also taken into consideration in order to better answer the main question of collocation knowledge among Swedish high school students. As mentioned in the method description, the target collocations were chosen from COCA with different frequencies ranging from 15 938 to 11. Frequency is defined as the number of occurrences in COCA corpus, e.g. *pay attention* appears 15 938 times in the database.

The collocations were classified into two frequency levels, namely high and low respectively, to simplify the analysis. The top half of the collocations with a higher number of appearances in COCA was categorised as high-frequent whereas the other half was low-frequent.

The data confirmed the hypothesis that students have better knowledge of high frequency collocations, with a mean score of 86%, whereas the mean score among the low frequency collocations was relatively lower (60%). Nevertheless, this is still high as earlier

research has suggested 56.6% to be a satisfactory level (González Fernández and Schmitt, 2015). Table 5 displays the division between high and low frequent collocations.
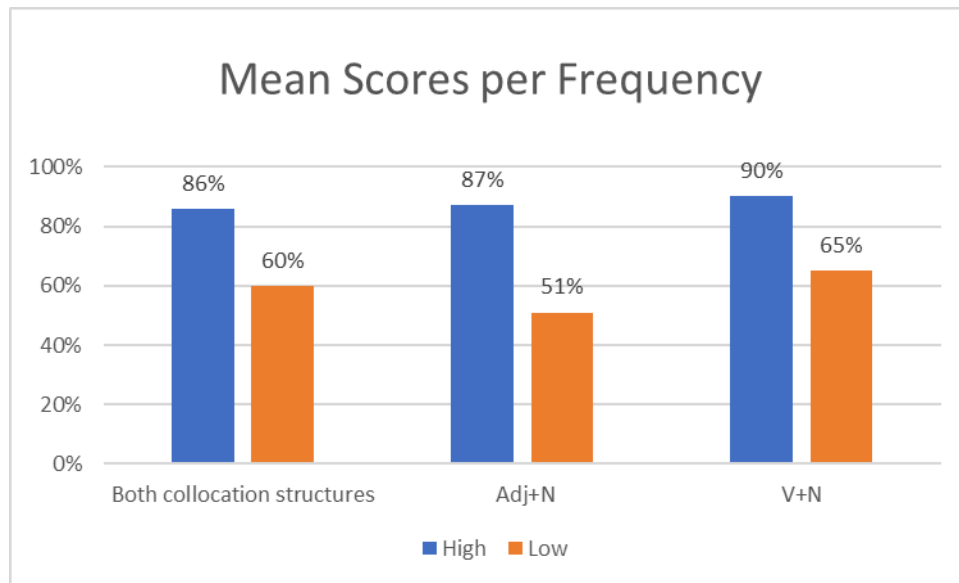
Table 5 demonstrates the differences in mean scores of the high respective low-frequency, within each structure and all collocations combined.

This confirms the hypothesis that L2 learners of English know high-frequent collocations more than low-frequent ones. Additionally, an analysis separating the V+N and Adj+N combinations also supports this claim, since the participants scored higher on the high-frequent collocations.

These findings also contrast the claim that learners of English seemingly do not acquire as many low-frequent collocations (González Fernández & Schmitt, 2015; Snoder, 2019); Although the receptive knowledge is lower among low-frequent combinations, the mean score percentage is higher than in previous research (González Fernández & Schmitt, 2015).

4.4. Other findings

After answering the main research question, there were other interesting patterns worth mentioning. Besides the aspects of frequency and the structures of V+N and Adj+N, gender distribution also appeared to be a curious variable. The girls' performances were shown to be superior to that of the boys and non-binary in this study. The finding contradicts the claim that boys have a greater level of vocabulary than girls (Sundqvist, 2010). However, as presented in Table 6, all genders received a mean score higher than 50% which suggests strong

collocational competence. A possible explanation for this finding may be that the girls spend more time on productive extramural activities such as reading, social media and surfing the internet.

Table 6 illustrates differences in collocation knowledge grouped by gender.

Another pattern worth mentioning is the correlation between receptive collocation knowledge and amount of hours exposed to English per week. The students were instructed to write how many hours they use English per week as well as estimate their skills in speaking, listening, reading and writing in English on a scale of 1-10. Most students answered 8 out of 10 on all four skills. Interestingly, reading and writing predicted receptive collocational competence more than speaking and listening. See table 7, 8, 9 and 10 below.

Tables 7 and 8 visualise the participants' estimation of their reading and writing ability in correlation with their total scores.

Tables 9 and 10 visualise the participants' estimation of their reading and writing ability in correlation with their total scores.

There were some cases where students had received a mean score higher than 50% despite estimating themselves lower than 5 out of 10 on listening and speaking. There was also a slight correlation between hours per week using English and receptive collocation knowledge, as can be seen in Table 11.
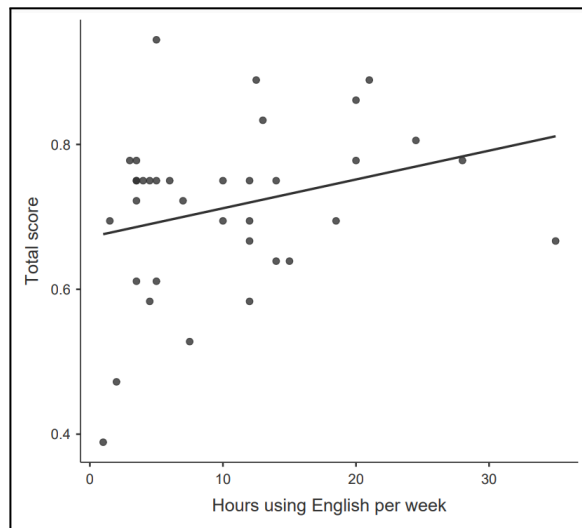

Table 11 shows the correlation between collocation knowledge and hours of using English per week.

5. Conclusions and further research suggestions

The purpose of this study was to investigate receptive collocation knowledge among Swedish high school students who are L2 learners of English. The results demonstrate an advanced

level of collocational competence among Swedish high school pupils, since the general mean score was 73.1 %. The findings also support the claim that L2 learners know more V+N collocations compared to Adj+N. The study confirms that L2 learners of English tend to recognise high-frequent collocations more than low-frequent ones. The data also suggest that reading and writing predicted collocation knowledge better than speaking and listening.

It was concluded that the level of receptive knowledge among high school students is high. However, it has been stated, as mentioned in the background, that recognition in lieu of production of collocation has been demonstrated to be less of a challenge for learners. Productive collocation knowledge is also suggested to be lower than receptive collocation knowledge. Therefore, testing Swedish high school students' productive competence can be a subject for further research.

In addition, measuring learners' recognition of collocations does not necessarily reveal their knowledge of the L1 meaning. The participants may have demonstrated that they recognise a correct collocation, but they may not have known the meaning of the collocation in their mother tongue (Nguyen & Webb, 2017, p. 316). Therefore, an interesting aspect for future research could be investigating the relation between English collocations with form and meaning in Swedish.

Gyllstad (2007) claims that a vast amount of exposure to English can boost collocation knowledge, extracurricular as inside the classroom. Also, according to Granger and Meunier (2008), teachers have to generalise and deprioritize vocabulary and phraseology due to lack of time. Therefore, it would be interesting to investigate further the relationship between extracurricular exposure and (receptive) collocation knowledge. Moreover, one can examine the relationship between specific activities (reading, writing, speaking and listening) and collocation knowledge.

References

Barfield, A., & Gyllstad, H. (Eds.). (2009). *Researching collocations in another language: Multiple interpretations.* Pelgrave and McMillian.

Barfield, A. (2013). Lexical collocations. *The Encyclopedia of Applied Linguistics*. Retrieved 2023-10-24 from https://doi-org.ludwig.lub.lu.se/10.1002/9781405198431.wbeal0690

Begagić, M. (2014). English language students' productive and receptive knowledge of collocations. *Explorations in English Language and Linguistics*, *2*(1), 46–67. Retrieved 2023-11-14 from https://doi-org.ludwig.lub.lu.se/10.1515/exell-2016-0003

Cloze test in English: Expert tips to solve accurately with solved examples. (2021, May 29). Retrieved 2023-11-06 from https://testbook.com/learn/english-grammar-cloze-test/

Davies, Mark. (2008-) *The Corpus of Contemporary American English (COCA)*. Available online at https://www.english-corpora.org/coca/. Retrieved 2023-12-06

Erickson, G. (2004). *Engelska i åtta europeiska länder: En undersökning av ungdomars kunskaper och uppfattningar* [English in eight European countries: A survey of young people's knowledge and perceptions]. Skolverket.

González Fernández, B., & Schmitt, N. (2015). How much collocation knowledge do L2 learners have? The effects of frequency and amount of exposure. *International Journal of Applied Linguistics*, *166* (1), 94 - 126. Retrieved 2023-10-19 from https://doi.org/10.1075/itl.166.1.03fer

Greenbaum, S. (1974). Some verb-intensifier collocations in American and British English. *American Speech*, *49*(1/2), 79–89. Retrieved 2023-12-24 from https://doi.org/10.2307/3087920

Gyllstad, H. (2007). *Testing English collocations: Developing receptive tests for use with advanced Swedish learners* [Doctoral Dissertation, Lund University].

Gyllstad, H. (2009). Designing and evaluating tests of receptive collocation knowledge: COLLEX and COLLMATCH. In Barfield, A & Gyllstad, H. (Eds.), *Researching collocations in another language: Multiple interpretations* (pp. 153–170). New York: Palgrave Macmillan.

Gyllstad, H. (2020). Measuring knowledge of multiword items. In Webb, S. (Ed.) *The Routledge Handbook of Vocabulary Studies.* (p. 387- 401). London: Routledge.

Gyllstad, H,. & Wolter, B. (2016). Collocational processing in light of phraseological continuum model: Does semantic transparency matter?. *Language Learning, 66*(2), 296–323. Retrieved 2023-09-17 from https://doi.org/10.1111/lang.12143

Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, *19*(1), 24–44. Retrieved 2023-11-21 from
https://doi-org.ludwig.lub.lu.se/10.1093/applin/19.1.24

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning, 61*(2), 647-672. Retrieved 2023-12-22 from https://doi.org/10.1111/j.1467-9922.2010.00621.x

McCarthy, M., & O'dell, F. (2017). *English collocations in use: How words work together for fluent and natural English: Self-study and classroom use.* (2nd Ed.). Cambridge University Press.

Masrai, A. (2023). Lexical knowledge and L2 general language proficiency: Collocational competence and vocabulary size as determinants of lexical knowledge. *Cogn Process, 24*, 289–300. Retrieved 2023-12-19 from
https://doi-org.ludwig.lub.lu.se/10.1007/s10339-022-01120-2

Meunier, S., Granger, F. (Eds.). (2008). *Phraseology: An interdisciplinary perspective.* John Benjamins Publishing Company.

Meunier, S,. Granger, F. (Eds.). (2009). *Phraseology: In foreign language learning and teaching.* John Benjamins Publishing Company.

Milton, J. (2009). *Measuring second language vocabulary acquisition.* Multilingual Matters. Retrieved 2023-12-23 from
https://ebookcentral.proquest.com/lib/lund/reader.action?docID=543908&ppg=5

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, *24*(2), 223–242. Retrieved 2023-10-26 from https://doi-org.ludwig.lub.lu.se/10.1093/applin/24.2.223

Nesselhauf, N. (2005). *Collocations in a learner corpus.* John Benjamins Publishing Company.

Nguyen, T. M. H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research, 21*(3), 298–320. Retrieved 2023-10-21 from https://doi.org/10.1177/1362168816639619

Olsson, E. (2016). *On the impact of extramural English and CLIL on productive vocabulary* [Doctoral Dissertation, Gothenburg University]. Retrieved 2023-12-02 from
http://hdl.handle.net/2077/41359

Research guides: Organizing your social sciences research paper: Quantitative Methods. (n.d.). Retrieved 2023-11-03 from
https://libguides.usc.edu/writingguide/quantitative#:~:text=Definition,statistical%20data%20using%20computational%20techniques.

Saudin, S., Sulyaningsih, I., & Meilinda, L. (2017). The investigation of productive and receptive competence in v+n and adj+n collocations among Indonesian EFL learners. *Indonesian Journal of Applied Linguistics*, *7*(1), 189–200. Retrieved 2023-11-13 from https://doi-org.ludwig.lub.lu.se/10.17509/ijal.v7i1.6874

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press.

Skolverket. (2011). *Ämnesplaner i gymnasieskolan på engelska* [Syllabi for upper secondary school in English]. Retrieved 2023-12-19 from https://www.skolverket.se/download/18.4fc05a3f164131a74181056/1535372297288/English-swedish-school.pdf

Snoder, P. (2019). *L2 instruction and collocation learning: Classroom intervention research on input processing with L1 Swedish adolescent learners of English.* [Doctoral Dissertation, Stockholm University]. Retrieved 2023-11-18 from https://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-166045

Sundqvist, P. (2010). *Extramural engelska : En möjlig väg till studieframgång* [Extramural English: A possible road to academic success]. *Kapet, 1*(1), 94–109. Retrieved 2023-12-08 from https://urn.kb.se/resolve?urn=urn:nbn:se:kau:diva-7087

Sylvén, L. K. (2018). *Internationell jämförelse av engelskspråkig undervisning i fem länder: Finland, Nederländerna, Tyskland, Spanien och Sverige* [An international comparison of English-language teaching in five countries: Finland, Netherlands, Germany, Spain and Sweden]. Skolverket. Retrieved 2023-11-27 from https://www.skolverket.se/publikationer?id=3893

Wang, Y., & Shaw, P. (2008). Transfer and universality: Collocation use in advanced Chinese and Swedish learner English. *ICAME Journal*, *32*, 201–232. Retrieved 2023-11-11 from https://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-14753

Webb, S., & Kagimoto, E. (2011). Learning collocations: Do the number of collocates, position of the node word, and synonymy affect learning? *Applied Linguistics, 32*(3), 259-276. Retrieved 2023-11-16 from https://doi-org.ludwig.lub.lu.se/10.1093/applin/amq051

Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, *63*(1), 91-120. Retrieved 2023-10-04 from https://doi-org.ludwig.lub.lu.se/10.1111/j.1467-9922.2012.00729.x

Appendix A

Frequency table of target collocations and distractors.

Collocations	verbatim phrase	lemma phrase	w1 lemma	w2 lemma	w1 form	w2 form
light frost	24	38	62687	3340	53929	3025
higher education	14171	14319	567720	231036	139409	230247
strong argument	331	739	193876	99403	150122	69218
happy birthday	6814	6848	170573	42727	158280	40376
heavy rain	2445	2555	78327	50093	69874	46813
bright colour	619	2277	59449	160011	49959	121051
single parent	1276	2236	148796	258181	148796	43802
healthy eating	1680	1105	69023	16142	60736	16136
severe pain	921	992	33278	101918	33089	95733
spectacular view	221	471	12709	182417	12709	133118
close relatives	891	1519	139704	26146	103513	19537
final draft	562	642	105688	30062	105688	27286
sole survivor	239	263	13543	24240	13519	9237
vivid imagination	194	217	7774	23622	7774	21745
decent meal	141	166	22751	48158	22712	29437
slight hesitation	69	126	23283	4491	17013	4256
tranquil countryside	11	11	1573	7352	1573	7283
brief summary	397	442	33229	20407	32399	18804

pay attention	15938	32796	365255	145669	192521	144742
waste time	10256	17446	32952	2018725	15192	1699431
play games	6684	42064	478740	445149	200898	135520
take a picture	3462	11055	1768822	165703	858510	103060
watch TV	4062	10373	318772	114859	135170	114808
raise children	2046	8039	190353	685426	63322	441998
review books	987	5138	48123	435387	19193	134252
cast a shadow	938	2629	41004	45022	31990	27497
follow a path	1351	3236	300266	72335	105100	60385
fight crime	491	959	154643	112978	70030	83017
take an exam	493	501	1768822	15933	858510	10610
save money	8113	12412	161469	437583	94434	436487
cause damage	2260	6233	170629	60384	63095	51252
hold a grudge	303	784	413837	2333	153835	1705
swallow pride	212	344	18013	24923	7609	24838
throw a tantrum	138	681	152388	2681	55237	1478
grant custody	32	168	47729	16070	11499	16070
gather evidence	432	1113	57562	162050	20847	161355

Distractors					
near relatives	12	145	112214	26146	112214	19537
concise summary	41	47	2287	20407	2287	18804
finished draft	3	4	5283	30062	5283	27286
lively imagination	13	14	7302	23622	6796	21745
small summary	4	4	382563	20407	316939	18804
soft frost	1	1	51824	3340	47482	3025
mild frost	1	6	14741	3340	13291	3025
upper education	2	2	37419	231036	37419	230247
advanced education	123	125	27563	231036	27563	230247
weak hesitation	0	0	45461	4491	36412	4256
faint hesitation	1	1	11768	4491	9747	4256
proper meal	58	77	41810	48158	41810	29437
adequate meal	2	6	17887	48158	17887	29437
gentle countryside	4	4	17606	7352	15856	7283
calm countryside	1	1	22255	7352	18356	7283
single survivor	19	21	148796	24240	148796	9237
lone survivor	165	177	9688	24240	9688	9237
thick rain	12	13	41420	50093	37314	46813

Distractors					
dense rain	14	14	10167	50093	10167	46813
tight relatives	0	0	25214	26146	20405	19537
sole parent	23	49	13543	258181	13519	43802
lone parent	13	25	9688	258181	9688	43802
steady argument	0	0	22152	99403	21739	69218
solid argument	54	71	43031	99403	43031	69218
last draft	29	31	132	30062	60	27286
harsh pain	4	6	17815	101918	15479	95733
relentless pain	7	7	5863	101918	5863	95733
dazzling view	9	19	3389	182417	3389	133118
dramatic view	31	62	33558	182417	30558	133118
blazing colour	0	0	1252	160011	1252	121051
glistening colour	0	0	1197	160011	1197	121051
healthful eating	64	64	1694	16142	1694	16136
hearty eating	4	4	2707	16142	2536	16136
merry birthday	0	0	9493	42727	9097	40376
cheerful birthday	0	0	4762	42727	4762	40376
give custody	12	0	1048189	16070	462452	16070
permit custody	0	0	33083	16070	12122	16070
show attention	1	9	536889	145669	219706	144742

Distractors						
give attention	97	0	1048189	145669	462452	144742
have a tantrum	0	0	10514314	2681	5023456	1478
start a tantrum	2	3	578246	2681	212064	1478
provide children	7	0	351941	685426	160610	441998
grow children	4	0	248803	685426	76073	441998
do games	0	0	8186412	445149	4501047	135520
set games	2	0	293144	445149	238088	135520
trail a path	0	1	13296	72335	3293	60385
seek a path	20	45	128234	72335	44438	60385
generate damage	3	5	51796	60384	21146	51252
do damage	0	0	8186412	60384	4501047	51252
drain time	2	14	14944	2018725	7285	1699431
lavish time	3	5	1028	2018725	244	1699431
rich imagination	13	15	89706	23622	78022	21745
write an exam	8	43	439865	15933	115267	10610
do an exam	0	0	8186412	15933	4501047	10610
snag a picture	0	4	3442	165703	1302	103060
snatch a picture	1	10	7383	165703	1909	103060
carry a grudge	26	59	168820	2333	65441	1705
scan TV	0	2	15695	114859	4890	114808
see TV	0	0	1958700	114859	1255990	114808

Distractors						
drop a shadow	0	5	134220	45022	46918	27497
throw a shadow	10	156	152388	45022	55237	27497
devour pride	0	0	4572	24923	1558	24838
absorb pride	0	0	19353	24923	7937	24838

Appendix B

Questionnaire in English.

Testing English Collocations!

My name is Louise Frifelt and I am writing my bachelor thesis in English at Lund University. The purpose of this study is to investigate Swedish high school students' knowledge and recognition of English collocations, which are natural combinations of words.

This survey will take 5-10 minutes of your time. Each question will have three options and you will pick ONE answer that you think is correct.

Participation is voluntary and you can refuse to participate anytime. There will be no penalty for withdrawing from the study. This survey is anonymous and all information will be kept confidential within this essay.Please answer no later than 22 December.

If you have any questions about this essay, please contact me on lo5247fr-s@student.lu.se

☐ I hereby consent to participate in this study

Age:*

Your answer

Gender:*

☐ Man
☐ Woman
☐ Non-binary
☐ Prefer not to say

Native language*

Your answer

On a scale of 1-10, to what level would you estimate your skills in **speaking** English?*

	1	2	3	4	5	6	7	8	9	10	
Poor	○	○	○	○	○	○	○	○	○	○	Excellent

On a scale of 1-10, to what level would you estimate your skills in **listening** to English?*

	1	2	3	4	5	6	7	8	9	10	
Poor	○	○	○	○	○	○	○	○	○	○	Excellent

On a scale of 1-10, to what level would you estimate your skills in **reading** English?*

	1	2	3	4	5	6	7	8	9	10	
Poor	○	○	○	○	○	○	○	○	○	○	Excellent

On a scale of 1-10, to what level would you estimate your skills in **writing** English?*

	1	2	3	4	5	6	7	8	9	10	
Poor	○	○	○	○	○	○	○	○	○	○	Excellent

How many hours per week do you use English?*

Your answer

Collocations are word combinations that appear together naturally. Your task is to pick ONE alternative that you think is the most natural combination of words in English. Here are some practice questions first!*

☐ achieve progress

☐ do progress

☐ make progress

Practise question 2.*

☐ large surprise

☐ big surprise

☐ huge surprise

Practice is over. Here are the questions from the questionnaire. Good luck!

1.*
- ☐ light frost
- ☐ soft frost
- ☐ mild frost

2.*
- ☐ keep silent
- ☐ stay silent
- ☐ remain silent

3. *
- ☐ make a family
- ☐ have a family
- ☐ start a family

4.*
- ☐ clever decision
- ☐ wise decision
- ☐ smart decision

5.*
- ☐ upper education
- ☐ advanced education
- ☐ higher education

6.*
- ☐ grant custody
- ☐ permit custody
- ☐ give custody

7.*

- ☐ weak hesitation
- ☐ slight hesitation
- ☐ faint hesitation

8. *

- ☐ show attention
- ☐ give attention
- ☐ pay attention

9.*

- ☐ adequate meal
- ☐ decent meal
- ☐ proper meal

10.*

- ☐ tranquil countryside
- ☐ gentle countryside
- ☐ calm countryside

11. *

- ☐ have a tantrum
- ☐ throw a tantrum
- ☐ start a tantrum

12. *

- ☐ lone survivor
- ☐ single survivor
- ☐ sole survivor

13.*
- ☐ heavy rain
- ☐ dense rain
- ☐ thick rain

14.*
- ☐ grow children
- ☐ provide children
- ☐ raise children

15.*
- ☐ tight relatives
- ☐ near relatives
- ☐ close relatives

16.*
- ☐ sole parent
- ☐ lone parent
- ☐ single parent

17.*
- ☐ do games
- ☐ play games
- ☐ set games

18.*
- ☐ seek a path
- ☐ follow a path
- ☐ trail a path

19.*

- [ ] brief summary
- [ ] concise summary
- [ ] small summary

20.*

- [ ] cause damage
- [ ] generate damage
- [ ] do damage

21.*

- [ ] drain time
- [ ] lavish time
- [ ] waste time

22.*

- [ ] lively imagination
- [ ] vivid imagination
- [ ] rich imagination

23.*

- [ ] analyse books
- [ ] criticise books
- [ ] review books

24.*

- [ ] take an exam
- [ ] write an exam
- [ ] do an exam

25.*
- ☐ assemble evidence
- ☐ draw evidence
- ☐ gather evidence

26.*
- ☐ solid argument
- ☐ steady argument
- ☐ strong argument

27. *
- ☐ fight crime
- ☐ oppose crime
- ☐ resist crime

28.*
- ☐ last draft
- ☐ final draft
- ☐ finished draft

29.*
- ☐ relentless pain
- ☐ severe pain
- ☐ harsh pain

30.*
- ☐ dramatic view
- ☐ spectacular view
- ☐ dazzling view

31.*

☐ take a picture

☐ snag a picture

☐ snatch a picture

32.*

☐ hold a grudge

☐ carry a grudge

☐ keep a grudge

33.*

☐ watch TV

☐ see TV

☐ scan TV

34.*

☐ blazing colour

☐ bright colour

☐ glistening colour

35.*

☐ healthy eating

☐ hearty eating

☐ healthful eating

36.*

☐ drop a shadow

☐ throw a shadow

☐ cast a shadow

37.*

☐ swallow pride

☐ devour pride

☐ absorb pride

38.*

☐ save money

☐ spare money

☐ preserve money

39.*

☐ cheerful birthday

☐ merry birthday

☐ happy birthday

40.*

☐ stunning smile

☐ dazzling smile

☐ bright smile

Thank you!