AT YOUR WITS' END: VOCABULARY LEARNING STRATEGIES IN DISCOVERING THE MEANING OF ENGLISH AND SWEDISH IDIOMS

Master's Thesis
Noomi Halonen
University of Jyväskylä
Department of Language and Communication Studies
English
June 2022

Tiedekunta – Faculty Humanistis-yhteiskuntatieteellinen	Laitos – Department Kieli- ja viestintätieteiden laitos
Tekijä – Author Noomi Halonen	
Työn nimi – Title At your Wits' End: Vocabulary Learning Strategies in Discovering the Meaning of English and Swedish Idioms	
Oppiaine – Subject Englannin kieli; (Ruotsin sivuainetutkielma)	Työn laji – Level Pro gradu -tutkielma
Aika – Month and year Kesäkuu 2022	Sivumäärä – Number of pages 156

Tiivistelmä – Abstract

Sanaston oppimiseen kuuluu olennaisesti yksittäisten sanojen lisäksi myös monisanaisten fraasien oppiminen. Idiomit ovat monisanaisia, merkitykseltään osin tai täysin läpinäkymättömiä ja ei-kirjaimellisia, melko muuttumattomia fraaseja, joiden merkitys ei ole yksittäisten sanojen merkitysten summa (Moon 1998: 5-8; Wood 2019: 32-33). Tutkimukset ovat osoittaneet, että idiomien erityispiirteet tuottavat usein haasteita kielten oppijoille (ks. esim. Mäntylä 2004; Szczepaniak 2006). Tässä tutkimuksessa selvitettiin vastaus seuraaviin tutkimuskysymyksiin: (1) Mitä strategioita lukiolaiset käyttävät englannin- ja ruotsinkielisten idiomien merkitysten selvittämisessä? (2) Miten hyvin merkitys pystytään selvittämään ja millaisia virheitä vastauksissa esiintyy? (3) Ilmeneekö eroja englannin- ja ruotsinkielisten idiomien selvittämisessä? Tutkimukseen osallistui yhdeksän lukion 2. vuosikurssin opiskelijaa. Osallistujien tehtävänä oli selvittää tuntemattomien englannin- ja ruotsinkielisten idiomien merkitys valitsemiaan keinoja käyttäen samalla ääneen ajatellen. Ääneen ajattelu ja tietokoneella tapahtuva toiminta tallennettiin ruutukaappausvideolla. Aineisto analysoitiin kvantitatiivisella ja kvalitatiivisella sisällönanalyysillä. Aineistosta löydettiin yhteensä yhdeksän erilaista strategialuokkaa, sekä neljä erilaista virhetyyppiä. Tutkimuksen tulokset osoittivat, että strategioista eniten käytettiin kaksikielistä sanakirjaa, arvausstrategiota, kuten kontekstista päättelyä sekä hakukoneen käyttöä (googlettamista). Lisäksi tuloksista ilmeni useita eroja englannin ja ruotsin välillä. Englanninkielisiä idiomeja selvitettiin useampia lyhyemmässä ajassa. Lisäksi lukiolaiset onnistuivat paremmin selvittämään englanninkielisten idiomien merkityksen, sekä olivat vähemmän taipuvaisia kirjaimellisiin käännöksiin, mutta erehtyivät yleistämään arvauksiaan kontekstista useammin. Myös strategioiden käytön määrässä sekä laadussa ilmeni huomattavia eroja. Nämä löydöt osoittavat, että idiomien merkitysten selvittäminen sekä niissä käytetyt strategiat vaihtelevat oppijasta ja kielestä riippuen. Tulosten perusteella olisi tärkeää, että vieraan kielen opetus antaisi oppilaille valmiuksia sanaston oppimisstrategioiden käyttöön sekä auttaisi ymmärtämään idiomien erityispiirteitä.

Asiasanat – Keywords

Vocabulary learning, Vocabulary learning strategies, Multiword units, Idioms, Think-aloud

Säilytyspaikka – Depository: JYX

Contents

1. INTRODUCTION

Vocabulary forms the core of a language. Words are the prerequisite for understanding and communicating. In fact, all the other aspects of language, pragmatics, syntax, phonology, and morphology, are built on lexical knowledge (Webb & Nation 2017: 5). Evidence suggests that both receptive and productive language skills are closely associated with and can be predicted by vocabulary knowledge (Qian & Lin 2020: 67-77). Moreover, vocabulary knowledge is essential in learning any kind of content (Webb & Nation 2017: 6). In Finland, language learning and teaching predominantly centers around vocabulary items. Some scholars argue that long-lasting and regular language instruction is not enough for foreign language learners to master sufficient vocabulary in the language (Webb & Chang 2012; Webb & Nation 2017: 152). Even when high language proficiencies are reached, gaining new vocabulary, and reinforcing partially known words, may still be a goal for many motivated advanced learners (Enström 2013: 169) as well as for L1 speakers since language is in constant change. For these reasons, vocabulary learning is a subject area that deserves attention. In this thesis, I will investigate the vocabulary learning strategies of Finnish upper secondary students while they engage in an English and Swedish vocabulary learning task

Learning vocabulary in another language does not only consist of learning single words but also of learning strings of words, such as *freak out, long time no see*, and *of course*. Some of these have been referred to with numerous, more or less related terms, such as chunks, collocations, multiword items, fixed expressions, and idioms (Wray 2002: 8-9). Although many of these terms share features, they should not be used interchangeably since there are differences in how they are defined and what kind of categories they represent. Wray (2002: 9) prefers to use the term formulaic sequence, which she defines in her often-quoted definition as:

> "a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar."

To this definition, Wood (2019: 30) adds that formulaic language, as the umbrella term for the phenomenon, holistically represents a single meaning or function similarly to single words. In this thesis, the terms multiword unit (mwu) and formulaic sequence will be used as superordinate terms to refer to a heterogenous set of items that consist of more than one word.

Due to the great magnitude of formulaic language, I have narrowed my focus to idioms, such as *miss the boat*, which in this thesis are defined as semi-transparent or opaque, somewhat fixed multiword units whose meaning cannot be derived from the meaning of its component words (see for instance Moon 1998: 5-8; Wood 2019: 32-33). The idiom *kick the bucket*, is a prototypical example of an idiom.

Since the 1980s, the importance of formulaic language has been widely recognized by many scholars (e.g. Lewis 2012; Nattinger & De Carrico 1992; Pawley & Syder 1983; Schmitt 2004; Wood 2015; Wray 2002). Wood (2015: 3) states that formulaic language is a fundamental aspect of language and communication, since it is linked to so many aspects such as discourse, fluency, and cognitive processing. Additionally, formulaic language seems to be a common linguistic phenomenon; Erman and Warren (2000) report that on average around half of English spoken and written texts consist of prefabricated multi-word combinations. Swedish, is also rich in formulaic sequences (e.g. Bolander 1989; Prentice & Sköldberg 2013). Idioms, as a particular category of formulaic language, have low frequency in both English and Swedish (Grant 2006; Moon 1998). But although the likelihood of encountering individual idioms is low, learners will unavoidably come into contact with idioms as a category at some point. In addition, idioms represent two linguistic aspects that are common, namely formulaic language, and figurative language. For these reasons, idioms are a useful object of research.

Many studies have shown that foreign and second language learners' knowledge of idioms is low (e.g. Macis & Schmitt 2017; Mäntylä 2004), and that they struggle with comprehending idioms (e.g. Szczepaniak 2006; Katsarou 2010). This is no surprise since idioms are morphologically, syntactically, and semantically complex units. The difficulties encountered by learners when learning idioms can be attributed to various factors, such as inability to recognize idioms as multiword units, non-literal meaning and/or figurative meaning (Cieślicka 2006; Kecskes 2000), and formal and semantic differences between L1 and L2. Consequently, it is essential to look at what strategic actions learners take to overcome these learning difficulties that are encountered while learning English and Swedish idioms.

In Finland, both English and Swedish have a special status, English as the international Lingua Franca, and Swedish as the second national language, also an obligatory subject in the curriculum for Finnish L1-speakers. The Finnish National Core Curriculum for upper secondary education (2019: 19, 59, 129, 176) emphasizes lifelong learning and the ability to continue language learning after finishing upper secondary studies. When learners possess skills in vocabulary learning strategies, learners will be better equipped for this task. According to the general learning objectives for Swedish and foreign languages (e.g. English), students

should be able to use language learning strategies efficiently (Lukion opetussuunnitelman perusteet 2019: 129, 176). Hence, this study provides new insights into what strategies Finnish upper secondary students actually use.

In its simplest form, vocabulary learning strategies (VLS) can be seen as actions that enhance vocabulary learning. For instance, a learner can read a novel and meanwhile guess meanings of unfamiliar words or mwus or look up their meanings in a dictionary. Webb and Nation (2017: 152) assert that VLS support autonomous vocabulary learning outside the classroom, which is crucial for developing a vocabulary "to the level that is required to understand the target language without support". Extensive research has shown that various VLS are linked to high language proficiency (e.g. Fan 2003, Gu & Johnson 1996).

Despite the large volume of studies focusing separately on various types of formulaic language and vocabulary strategies, much less is known about the vocabulary learning strategies for learning mwus and idioms (Wood 2015: 167, Gu 2019: 281). Only recently, more studies focusing on particular VLS, such as dictionary use, and mwus have emerged. Consequently, Gu (2019: 281) makes an urgent call for more research effort on strategies used for learning mwus. Furthermore, less attention has been paid to receptive skills, such as how language learners handle formulaic sequences when compared with the number of studies investigating productive skills of speaking and writing (Wood 2015: 172). Vocabulary research has also predominantly concentrated on quantitative studies, which is why Webb (2019a: 7) petitions for more qualitative studies. Lastly, no studies on vocabulary learning strategies conducted in the Finnish context could be found, let alone studies in which English and Swedish idiom comprehension, and strategy use was compared. Therefore, with this thesis, I aim at trying to fill these gaps in research.

In this study, data was collected from nine Finnish second-year upper secondary students. The data consisted of video screen capture recordings that also included think-aloud protocols. With this study, my aim is to contribute to the knowledge about what on-line vocabulary learning strategies are used to discover the meaning of unfamiliar English and Swedish idioms. Additionally, I will investigate how well Finnish upper secondary students are able to discover the meaning of English and Swedish idioms, and what types of errors they make. Lastly, I will compare the participants' task performances with regard to English and Swedish idioms. In this thesis, I aim at answering the following research questions:

1. What vocabulary learning strategies do Finnish upper secondary students use in discovering the meaning of English and Swedish idioms while thinking aloud?

2. How well are students able to discover the meaning of English and Swedish idioms? What types of errors do they make?
3. Are there any differences between discovering the meaning English and Swedish idioms?

Next, I will present some key terminology and theory about vocabulary knowledge and vocabulary learning, and then move on to a more in-depth presentation of idioms. After this, in chapter 5, I will give an overview of vocabulary learning strategy theory and research which will be followed by a more thorough exploration of some strategies in section 6. In section 7, I will present the data collection methods and the method of analysis. Then, in sections 8, 9 and 10 I am going to report the quantitative and qualitative results of this study. Lastly, in section 9, I will discuss these results and their significance in light of previous research.

2.VOCABULARY KNOWLEDGE

Before proceeding to examine idioms and vocabulary strategies, it is important to develop an understanding of vocabulary knowledge, since it is essential for being able to understand theoretical perspectives of vocabulary and current vocabulary research. The next two sections will focus on the several dimensions of vocabulary knowledge (2.1), and the several different aspects that are included in knowing multiword units (2.2).

2.1 Aspects of vocabulary knowledge

There are many aspects to know about a particular word (Nation 2013: 44), such as spelling, pronunciation, derivations, collocations, semantic associations, and grammatical functions (Webb 2013; Webb & Yanagisawa 2019). This view of vocabulary knowledge as a multifaceted construct has been recognized by many scholars (e.g. Ellis 1995; Nagy & Scott 2000; Nation 2013; 2019; Richards 1976). Words are also known to different degrees (Nation 2013: 44), that is, "different levels of strength and detail, and different levels of fluency" (Nation 2019: 15). This accords with Henriksen's (1999) partial-precise vocabulary dimension. So, instead of being a clear-cut matter, vocabulary learning is an incremental process that proceeds gradually from partial knowledge of a word to a fuller account of its several aspects. As a result, the specific aspects and stages of vocabulary knowledge that are focused on should be made explicit in vocabulary research since the point at which a word/mwu is known and what counts as knowledge varies considerably.

The distinction between receptive and productive vocabulary knowledge is a generally accepted dimension in vocabulary research (Henriksen 1999). For Nation (2013: 46-47) receptive refers to comprehending input through listening and reading, whereas productive refers to producing language in speech or writing to convey a message. These definitions can nevertheless, be misleading since meaning is also produced during listening or reading (Nation 2013: 47). However, the terms are more suitable than the previously used passive and active vocabulary, due to the passive connotation associated with reception (Nation 2013: 47). Today, it is more common that the type of vocabulary test is described with the terms recognition (receptive), and retention (productive). Up to now, several studies have confirmed that learners' receptive vocabulary is larger than his or her productive vocabulary and that the gap between receptive and productive vocabulary is smaller for high-frequency words, but it tends to

increase as the frequency of words decreases (Laufer & Paribakht 1998; Shin et.al. 2011, Waring 1997; Webb 2008a). This means that low-frequency words, such as idioms, are more likely to be known receptively than productively. However, the learners may also have partial knowledge of the word (Webb 2008a).

Vocabulary studies usually distinguish between receptive and productive vocabulary knowledge. Although learning effects of the vocabulary task are not measured in this study, the idiom meaning discovery task could be categorized as receptive. Therefore, it is more likely to result in higher receptive gains given that the other learning conditions favor learning. Since learning idioms is already complicated as such, receptive learning, being less difficult than productive learning (Ellis & Beaton 1993; Mondria & Wiersma 2004: 86, 97), is an adequate goal particularly with intermediate and lower-advanced learners of English and Swedish. Moreover, if productive learning follows receptive (Melka 1997; Henriksen 1999), there is value in receptive tasks.

In vocabulary research, there is also a distinction between vocabulary breadth, i.e. the number of words known, and vocabulary depth, i.e. how well a word is known (Webb 2019a: 6). This distinction affects the type of test that is used to measure vocabulary knowledge. Whereas tests of vocabulary breadth often only test receptive knowledge of the form-meaning link (e.g. Schmitt et.al. 2001), tests of vocabulary depth test multiple aspects of knowledge (e.g. Read 1993). Anderson and Freeboy (1981) were the first ones to define vocabulary depth. In their definition, depth is juxtaposed with all the distinctions an adult native speaker understands about a word (Anderson & Freeboy 1981: 93). The problem with this definition, as pointed out by Read (2004: 213), lies in that there presumably is a lot of variability in how vague, polysemous, and technical words are understood by native speakers.

2.2 Knowing a multiword unit

The most comprehensive and widely recognized description of what is included in knowing a word is Nation's (2013) categorization. It represents the components approach of depth since it describes the separate elements of vocabulary knowledge (Webb & Yanagisawa 2019: 375). The original categorization only covered single words but a later adaptation developed by Nation and Webb (2011: 190) extended it to multiword units since both can be approached in similar ways (Webb & Nation 2011: 189). In this adapted model, knowing a multiword unit

involves knowing receptive (R) and productive (P) aspects of form, meaning, and use (Nation 2013: 48). The categorization is presented in table 1 below:

Table 1. What is involved in knowing a multiword unit (MWU), adapted from Nation and Webb (2011: 190)

Form	spoken	R P	What does the MWU sound like? How is the MWU pronounced?
	written	R P	What does the MWU look like? How is the MWU written and spelled?
	word parts	R P	What words are recognizable in this MWU? What words are needed to express the meaning?
Meaning	form and meaning	R P	What meaning does this MWU signal? What MWU can be used to express this meaning?
	concepts and referents	R P	What is included in the concept? What items can the concept refer to?
	associations	R P	What other words or MWUs does this make us think of? What other words or MWUs could we use instead of this one?
Use	grammatical functions	R P	In what patterns does the MWU occur? In what patterns must we use this MWU?
	collocations	R P	What words, MWUs, or types of MWUs occur with this one? What words, MWUs, or types of MWUs must we use with this one?
	constraints on use (register, frequency etc.)	R P	Where, when, and how often would we expect to meet this MWU? Where, when, and how often can we use this MWU?

The model captures well the complex and multidimensional nature of vocabulary. It can be applied to both English and Swedish since formulaic language is a distinct characteristic of

both languages. The model is not a hierarchy but rather a description, yet Nation (2019: 15) notes that not all of the aspects are equally important. In general, the form-meaning connection is regarded as the most important stage in learning a new word/mwu, which is the aspect focused on in this study. Additionally, receptive aspects of form and meaning are also relevant. These include the written form, word parts, concepts and referents, and associations. Next, I will discuss these aspects from a learning point of view, and in relation to the characteristics of multiword units.

When learning single words, words can be decomposed into smaller parts, morphemes, which may be inflections or derivations such as suffixes or prefixes. Learning them would imply becoming aware of the word consisting of these parts and understanding their semantic contribution to the meaning of the word. Learning mwus is slightly different because it could be argued, based on the comparison of the original and the adapted model of Nation's categorization that not morphemes but single words are the minimal units that make up the mwu. As a result, mwus may be approached from a more holistic perspective while paying less attention to the lexical makeup of individual words. This is only a hypothesis and in reality the learner may tackle the double learning objective, that is, learn the words in the mwu and the word parts of the component words. In conclusion, the learning of mwus seems to be a more complex process than learning single words.

Making a connection between form and meaning is one of the primary steps in learning new words. Prior to making this connection, the learner may recognize the form and have knowledge of the concept, for instance via L1 or word knowledge, without yet making the connection (Nation 2019: 18). A form-meaning connection is easier to establish for loanwords and cognates, that is, words that have the same origin (Nation 2019: 19). In relation to mwus that are semantically transparent, the learner may infer the meaning of the mwu by adding up the meanings of single words (Nation 2019: 21). Regarding idioms, the learner may have to integrate a completely new collective meaning to a string of words that otherwise may look familiar.

Knowing the concept of a word entails having a deeper understanding of its meaning which in turn will allow understanding it in various contexts (Nation 2013: 50). Words can have a core meaning that "runs through all or most of their uses" (Nation 2019: 19). The different uses of a word are called senses. In addition to having a core meaning realized by various senses, single words, for example the word *bear*, can also be polysemous, that is, have multiple meanings. Mwus and idioms in turn have less senses of a core meaning and are less polysemous than single words. However, understanding the concept of the mwu may require

understanding in which particular senses or meanings the single words that the idiom consists of, are used. This is where the conceptual knowledge of single words comes into play. In the case of figurative idioms, their meaning can be seen as a figurative extension of the literal meaning of the word(s) for instance in the case of prepositions (Boers & Lindstromberg 2008: 21; 28). Additionally, learning the meanings of idioms may include learning culturally specific concepts that do not exist in the learners' L1.

The mwu can be associated with other related words or mwus in the lexical network. These can for instance be synonyms, opposites, or hyponyms (Nation 2019: 20). According to Nation (2019: 20), associations are often learned implicitly. Nevertheless, the intentional learning of associations may be encouraged by the use of thesaurus and dictionaries. For example, monolingual dictionaries may provide synonyms and other similar concepts as definitions for the mwu or idiom in question. To illustrate this, the Swedish idiom, *gå ur tiden* (literally: 'go out of time'), is defined in a monolingual dictionary with a synonymous word, *dö* ('die'), and a synonymous multiword unit, *ge upp andan* ('give up the ghost').

It is highly unlikely that all of the aspects described above would be learned simultaneously. Instead, the learner may during a long period of time, at various occasions, both intentionally and implicitly, acquire some, most, or all of these aspects. Another remark worth making is that vocabulary knowledge is not only affected by increasing knowledge of single words but also by learning different patterns and systems of the language (Nation 2013). For example, learning common affixes will facilitate learning words that have those affixes. As learners are able to generalize rules and notice patterns in language, the burden of learning individual aspects diminishes.

3. VOCABULARY LEARNING

In this chapter, I will present differences between intentional and incidental learning (3.1), list conditions that are required for vocabulary learning (3.2), introduce the four strands of learning multiword units (3.3), and finally, delve into learning multiword units (3.4).

3.1 Intentional and incidental learning

Vocabulary can be learned in a range of different ways and with a plethora of techniques. One reason for this is that language is present everywhere. Vocabulary learning research usually focuses on either intentional/deliberate or incidental learning. In applied linguistics, incidental learning is defined as learning that occurs as a by-product of a meaning-focused task (Webb 2019b: 225). Even if the term incidental alludes to a lack of intention, it is likely that learners have certain degrees of intention to learn (Webb 2019b: 225-226). This has been illustrated with an eye-tracking study which shows that L1 and L2 participants spent a longer time reading unknown words in a meaning-focused reading task (Pellicer-Sánchez 2016). The problematicity of the term "implicit" is also supported by the fact that the amount of intention is difficult to measure. In addition, attention and noticing are necessary for learning (Hulstijn 2003: 361; Nation & Webb 2017: 68). But whereas the degree of intention to learn new words may vary both in incidental and intentional learning, only in intentional learning is there a deliberate goal to learn new words (Hulstijn 2003: 360). To avoid the problematic question of intention, Webb (2019b: 226) advocates replacing the terms with meaning-focused learning and language-focused learning (see also 3.3).

It has been established that incidental learning of L2 words can occur through different modes of meaning-focused input, such as reading (e.g. Day et.al. 1991; Waring & Takaki 2003), listening (e.g. Pavia et.al. 2019; Zeeland & Schmitt 2013), and viewing (e.g. Peters & Webb 2018; Rodgers & Webb 2020). Studies on intentional learning are far more scattered into investigations of certain theories of learning, language instructions, memorizing techniques, and learning strategies (Hulstijn 2003; Lindstromberg 2019). Of the vast number of different types of intentional learning, Lindstromberg (2019) reviews semantic elaboration (for semantic elaboration with idioms see e.g. Boers et.al. 2004; Boers et.al. 2007), the elaborative processing of lexical form (see e.g. Barcroft 2015), and the keyword method (e.g. Ellis & Beaton 1993).

Other common intentional learning techniques include paired-associate learning with flashcards (e.g. Nation 2013; Nakata 2019), and list learning (e.g. Yamamoto 2014). Moreover, a long list of common L2 exercise types such as multiple-choice, fill in the blank, sentence production, and cloze-activities can be subsumed under intentional learning (Webb 2019a: 5; 2019b: 228). As can be seen from the task types above, intentional learning occurs often in a formal classroom setting, whereas implicit learning is more likely to take place outside the classroom.

Research findings show that intentional learning is more efficient than incidental learning (Laufer 2003; Webb 2019b: 231). In other words, intentional learning leads to better results in less time than implicit learning. This can be illustrated with Laufer and Girsai's (2008) study. They investigated 75 high school EFL-learners' learning of single words and collocations in three instructional conditions: meaning-focused instruction (MFI), non-contrastive form-focused instruction (FFI), and contrastive analysis + translation (CAT). The first condition fulfills the definition of implicit learning and the latter two conditions fit in the definition of intentional learning. The results confirmed the superiority of intentional learning; participants in the CAT-group and the FFI-group scored much higher on the active and passive recall tests than the MFI-group whose learning gains were minimal. But before drawing any conclusions, it has to be noted that the efficiency of learning depends on many other conditions as well (see 3.2).

Several researchers have argued that the results of incidental learning may be qualitatively different. While intentional learning may focus on the form-meaning connection, incidental learning may foster the learning of i.a. grammatical features and collocation (Nation & Webb 2011: 307; Webb & Nation 2017). Additionally, incidental learning may develop more gradually (Webb 2019b: 233). But unlike in intentional learning, for significant incidental learning gains to occur, massive amounts of input have to be encountered (Webb 2019b: 233) to ensure multiple encounters with the word/mwu, and around 95-98% of the words in input have to be known to the learner (Liu & Nation 1985; Hu & Nation 2000). For these reasons, both types of learning are needed, intentional learning for its effectiveness, and incidental learning for the gradual building of a rich and large vocabulary that could not intentionally be obtained (Webb 2019b).

The meaning discovery task in this study, on the one hand, leans more towards the intentional end of the continuum, since the explicit goal of the task was to discover the meaning of idioms. Firstly, the participants were aware that the study focuses on vocabulary. Even if the participants were not tested on meaning recognition or other aspects of learning, mere

awareness of learning may also trigger expectations of a test (Hulstijn 2003: 356). Secondly, the participants were instructed to discover the meaning of specifically chosen items (see also Barcroft 2015: 42). However, the participants had the freedom to choose how they approach the task and what strategies they use (see also Lindstromberg 2019: 241), as when learning in an incidental context. Moreover, the participants used the guessing from context strategy, which is characteristic for incidental learning, nevertheless in a more deliberate manner. As can be seen, categorizing a task as implicit or explicit is not straightforward. Additionally, it should be remembered that discovering the meaning is only a step towards learning a vocabulary item, and for this reason, the using the term learning may be misleading. Implicit learning and the factors affecting it will further be touched upon in section 6.2 in relation to guessing from context, and intentional learning strategies for idioms will be addressed in section 6.4.

3.2 Conditions for learning

Vocabulary learning is affected by certain learning conditions. Nation and Webb outline five different conditions that facilitate vocabulary learning: repetition, noticing, retrieval, varied encounters and varied use, and elaboration (Nation & Webb 2017: 61-76). Repetition, i.e. the number of encounters with the word, is a quantitative aspect, whereas the rest of the conditions are qualitative in nature. Noticing occurs when the learner pays attention to the word, for instance when using strategies to discover the meaning of the word. Retrieving the word occurs when the word is recalled either receptively or productively, for instance when using flashcards. Varied encounters and varied use refer to encountering and using the word in different contexts. Elaboration occurs when knowledge of a word is enriched, for instance by encountering more aspects of form, meaning, and use. The fulfillment of some or all of these conditions may enhance and strengthen the learning of vocabulary items, and vice versa, the nonfulfilment or lack of these conditions may weaken learning.

3.3 The four strands and multiword units

Nation (2013; see also Nation & Webb 2017: 182-187) addresses vocabulary learning from the perspective of four strands. The four strands include meaning-focused input, meaning-focused output, language-focused learning, and fluency development. The model can also be

implemented into learning multiword units (Nation 2013: 497-508). Meaning-focused input refers to reading or listening for comprehension, whereas meaning-focused output refers to speaking or writing with a communicative focus. Doing exercises that deliberately focus on multiword units is part of language-focused learning. The aim of fluency development is the automatization of already known vocabulary items. The four strands highlight the different modes and techniques of learning, but also the value of each strand for multifaceted vocabulary knowledge. In this study, the participants will engage in a task belonging to the language-focused learning strand, since language items, not meaning nor fluency, are the focus of the task. Usually, language-focused learning includes explicit teaching of vocabulary, nevertheless, intentional independent learning also fills its criteria.

3.4 Learning multiword units

There are several plausible reasons for learning formulaic sequences. Firstly, Pawley and Syder (1983) argue that knowledge of familiar lexicalized and institutionalized sentence stems account for the nativelike selection of words, i.e. ability to produce idiomatic language. Secondly, knowledge of formulaic sequences aids receptive fluency by rendering language more predictable (Boers 2019: 143). Thus, if ready-made chunks are used in encoding, more capacity is released to other functions, such as pronunciation (Pawley & Syder 1983: 208). Thirdly, the use of formulaic sequences contributes to speech fluency (Wood 2010) since they may allow the speaker to speak faster, produce longer runs of speech with fewer pauses in the middle of a clause or phrase (Wood 2015: 87-88; Wood 2010). This advantage can be explained with a proposition according to which formulaic sequences are retrieved from long-term memory as chunks, and as a result, the limit to how many words can be kept in short-term memory in sentence production can be bypassed (Wood 2015: 57). All in all, knowledge of formulaic sequences is a considerable advantage for L2 speakers, and hence, it is not a surprise that L2 speakers who use more formulaic language may seem more proficient (Boers et.al. 2006; Stengers et.al. 2011; Wood 2015: 140).

One of the advocates for the learning of mwus is Michael Lewis (2012), the developer of the lexical approach, a theoretical and pedagogical proposal for language learning. Lewis (2012) rejects the structuralist, grammar-based view of language learning according to which "grammar is creative while words are like building bricks, fixed packages of meaning" (Lewis 2012: 37). Instead he argues for basing learning not on words, but on lexical items, such as

sentence frames, collocations and institutional expressions, which can subsequently be used as resource material for learning grammar. In sum, language consists of "grammaticalised lexis" in Lewis' terms (Lewis 2012: 89). This is reasonable, since it allows the learners to surpas their current grammatical level by memorizing grammatically correct phrases, and directly using them in communication (Wood 2015: 78). With whole phrases as a starting point, Lewis (2012: 115-132) proceeds with giving several practical suggestions for teaching lexis. Distinctive to these exercises is their emphasis on the co-occurence of words. Furthermore, Lewis (2012: 98, 111-113) recognizes the prominent role that metaphor and underlying metaphorical patterning has in everyday language.

Boers (2019) reviews some of the factors that affect the learning of mwus. The relevance of these factors depend on the learner, the type of mwu in question, and the type of learning. As for mwus in general, one factor that Boers (2019: 145) argues is likely to affect learning concerns the familiarity with the mwu's component words, unfamiliar words being more challenging to recall than mwus consisting of familiar words. In incidental learning, the salience of the mwu in the text, and the amount of exposure (Webb et.al. 2013) play a significant role. The results of Bishop (2004) support the former, as he found that typographical salience increased the look-up of glosses, and consequently, improved comprehension of unknown mwus. In deliberate learning, the results depend on the instructional techniques and activities used. Moreover, interlexical similarity may affect the learning of mwus (Boers 2019: 152). Finnish is a synthetic language that uses inflections, as opposed to English and Swedish, which make more use of particles and prepositions. These structural differences may predict difficulties for Finnish learners when learning English and Swedish mwus such as phrasal verbs, since such a category of mwus is nonexistent in Finnish. How polysemous the constituent words are may also have an impact on comprehending mwus, since choosing a wrong sense for individual words may easily lead to misinterpretations. (Boers 2019: 145). In sum, learning mwus is affected by multiple factors. If other additional factors related to idiom learning (see 4.3) are considered, it can be hypothesized that learning mwus is a complex process.

Studies on English mwus and idioms abound, whereas there is far less Swedish literature on the topic. Due to the lack of research on Swedish idioms, some studies focusing more generally on mwus will be presented here. In Swedish literature, the following terms are often used: *flerordsenheter* ('multiword units'), *fraseologiska enheter* ('fraseological units'), *konventionaliserade fraser/uttryck* ('conventionalized phrases/expressions'), *figurativa ordförbindelser* ('figurative expressions'), and *idiomer* ('idioms'). The multiplicity of Swedish

terms is once again evidence of the heterogeneity of formulaic language. Studies on Swedish mwus and idioms have been conducted in the Swedish context with a focus on L2-learners' oral skills. Sweden of today, is a multilingual country, inhabited by speakers of, for example, Arabic, Finnish and Bosnian. Researchers have sought to compare L1-learners and L2-learners with regard to their knowledge of mwus and idioms, and to describe features of L2 language production.

Knowledge of formulaic sequences has been considered to be one of the yardsticks of high-level and near-native proficiency in a language (Hyltenstam 2016). Abrahamsson and Hyltenstam (2009) investigated the effect that age of onset, i.e age of arrival to Sweden has on perceived and actual (linguistically measured) Swedish nativelikeness. In their study, the speech of 195 L2-participants of various ages was judged by native speakers. After this, 41 participants that had been rated as native speakers completed ten linguistic tests that were further analyzed. In two of the tests, knowledge of idioms and proverbs was measured with oral fill-the-gap tasks (Abrahamsson & Hyltenstam 2009: 279). The results show that as the age of onset, i.e. the point of moving to Sweden increased, the knowledge of idioms decreased, a finding also obtained by Prentice (2010). Moreover, only 58% of the learners who have moved to Sweden between the age of 1-11 reached nativelike scores on idiom tests (Abrahamsson & Hyltenstam 2009: 287). The differences in early and late Swedish L2-learners were statistically significant (Abrahamsson & Hyltenstam 2009: 286). However, the least nativelike scores of the ten tests were on the proverb test. These results show that even near-native L2-learners lack in their knowledge of mwus. In conclusion, since idioms and proverbs seem to pose a challenge to L2-learners living in Sweden, they must be even more challenging for foreign language learners that likely are less exposured to Swedish input. Thus, the goal of nativelike command of idioms, even more so of proverbs, is unrealistic.

Other studies that have examined the knowledge of Swedish mwus have predominantly centered around the language production of Swedish L2-learners. It has been established that L1 speakers use more conventional expressions than L2 speakers (Ekberg 1997, 1998 as cited in Prentice & Sköldberg 2010: 12; Ekberg 2013; Hyltenstam 1992; Prentice & Sköldberg 2010). Prentice (2010: 5-6) defines conventional expressions as established more or less fixed forms for expressing certain content. Therefore, the term encompasses various types of fixed mwu, also idioms. Prentice and Sköldberg (2010) looked at the conventionality and institutionalization of Swedish figurative expressions in the student essays of 175 pupils with various linguistic profiles. They found three types of figurative word combinations which were categorized into conventional, partially modified, and novel word combinations. They found

that L2 speakers used plenty of figurative language, however, it was qualitatively different from L1 use. Whereas L1 pupils used more conventional expressions, L2 students used more partially modified conventional expressions, in which constituents were changed, switched, removed or added (Prentice & Sköldberg 2010: 24; see also Ekberg 2013). Hyltenstam 1992 (as cited in Prentice & Sköldberg 2010: 13) also observed how non-natives speakers deviate from the standard form. He found, in line with Ekberg (2013: 266) that native speakers used more approximations of standard forms, for instance when mixing prepositions. However, it has been observed in some studies that native speakers also appear to deviate from the norm, albeit to a lesser degree (Hyltenstam 1992, Ekberg 1997; 1998 as cited in Prentice & Sköldberg 2010). In sum, the oral production of L2 learners differs from L1 production both quantitatively and qualitatively. These findings indicate that mwus are difficult for L2 learners.

4. IDIOMS

In this section, definitions of idioms, categorizations of idioms, and studies on idioms, including the factors that affect learning idioms will be presented and discussed.

4.1 Definition

Formulaic sequences have been categorized into several subcategories that each have their own defining features (Wood 2015). However, the categories are problematic since there are considerable overlaps and imprecision, which can lead to different interpretations (Wood 2019: 30). Likewise, defining idioms is no easy task due to the enormous variability in definitions and a lack of general consensus (Grant & Bauer 2004: 44; Liu 2008). To make the case even more puzzling, the term idiom, together with the term collocation, has been used when referring to formulaic language in general (Wood 2015: 36). Definitions of idioms range from broad (Hockett 1958; Makkai 1972) to narrow (Grant & Bauer 2004). Various structural, semantic and syntactic criteria have been applied in order to define what counts as an idiom. Moreover, some scholars strictly demand the fulfillment of all criteria they have presented, whereas others are content with the fulfillment of only some of these criteria. Liu (2008) provides a comprehensive summary of some of the most prominent definitions of idioms. He sums that due to the complexity and ambiguity of idioms, presenting a single definition would not be adequate. Additionally, Wood (2019: 32-33) summarizes the past efforts to define idioms into five following criteria: 1) minimum length of two words, 2) semantic opacity, 3) non-compositionality, 4) mutual expectancy, and 5) lexicogrammatical invariability/ frozenness / fixedness. Most of these will be touched upon in what follows.

One matter of debate has been the number of items that an idiom is comprised of. At one extreme end, Hockett (1958: 172-173) counted morphemes whose meaning cannot be deduced from the structure, as idioms. Other scholars have counted single words as idioms (Katz & Postal 1963; 1973 as cited in Liu 2008; Makkai 1972), whereas some scholars have restricted idioms to multiword units (Weinreich's 1969 as cited in Liu 2008; Grant & Bauer 2004). This debate is understandable since the boundaries between single words and mwus can sometimes be blurry. For instance, Moon (1998. 8) notes that some mwus can have single word cognates (*break the ice* and *ice-breaker*) (Moon 1998: 8). Nevertheless, most scholars agree on

that idioms can be, or are multi-words, i.e. phrases, or clauses that consist of at least two words (Liu 2008).

Two semantic features, non-compositionality semantic opacity, are probably the most often cited criteria for defining idioms in literature (e.g. Fraser 1970; Hockett 1958; Makkai 1972; Moon 1998) and research (e.g. Boers et.al. 2007: 43; Macis & Schmitt 2017). Moon (1998: 8) explains non-compositionality as follows: "The meaning arising from the word-by-word interpretation of the string does not yield the institutionalized, accepted, unitary meaning of the string." In other words, summing the meanings of single words do not yield the meaning of the whole. For example, unlike the interpretation of *get a job* can worked out by adding up the meanings of its parts, the meaning of *break the ice*, which according to *Cambridge Dictionary* means "to make people who have not met before feel more relaxed with each other", cannot be computed or analyzed from the component words. Non-compositionality may arise from for instance, metaphor, grammatical violations, or unique lexis (Moon 1998: 5). According to Grant and Bauer (2004: 48), compositionality can be tested by replacing words in the mwu with their dictionary definitions.

A very similar criterion to non-compositionality is semantic opacity, which means that the meaning of the idiom is not transparent or literal. However, once again, there are differing views among scholars. For instance, Fernando (1978; 1996 as cited in Liu 2008) classified idioms into pure (non-literal), semi-literal, and literal idioms, and accepted all three categories as idioms. Others have regarded only semi-transparent or opaque items as idioms (Moon 1998: 5; Grant & Bauer 2004). However, idioms can have both a literal and a figurative meaning, and in fact, in Weinreich's (1969 as cited in Liu 2008) narrow definition, idioms are limited to only these types of lexical items. Some scholars, such as Grant and Bauer (2004) have been absolute regarding non-compositionality, by excluding all expressions, whose meaning hypothetically can be decoded, meanwhile Moon (1998: 8) and Kjellmer (1996: 82-83) interpret non-compositionality less strict. The conceptions of non-compositionality and transparency are problematic, since they are susceptible to subjective interpretation at a certain time (Moon 1998; Grant & Bauer), and consequently, two different people may rate the same idiom differently.

Some of the formulaic sequences have been termed frozen phrases or fixed expressions (Wray 2002: 9). Moon (1998) used the broader term FEIs (fixed expressions including idioms) in her corpus study. These labels, often used to describe idioms, imply that unlike single words prone to lexicogrammatical changes, idioms are completely, or to a certain extent lexicogrammatically invariable (Wood 2015: 44). This means that idioms allow only little, or

no morphological, lexical or syntactic variation. Within the framework of transformational grammar, Fraser (1970) composed a frozenness hierarchy that presents seven different levels of transformations. Whereas normal language permits unrestricted syntactic transformations, idioms range from being completely frozen, to permitting certain grammatical operations such as adjunction, insertion, and permutation. Therefore, idioms can be placed across various ranges of a spectrum. The proposition that not all idioms in fact are completely frozen has been confirmed in corpus studies (Grant 2006; Moon 1998). Like English idioms, Swedish idioms allow limited lexical and syntactic variations, such as when a word in an idiom is swapped to a synonym (*hård/svår nöt att knäcka*, 'a hard nut to crack') (Prentice & Sköldberg 2013: 202). Moreover, deviations of the fixed standard form, such as passivization, may occasionally result in literal interpretations and the loss of the figurative denotation (Enström 1996; 106; as cited in Prentice & Sköldberg 2013: 202-203).

A few more criteria for defining idioms commonly occur in literature. Firstly, many idioms are also metaphorical or figurative. Whereas Moon (1998: 4-5, 22-23) recognizes idioms as metaphorical expressions, and includes metaphors as one category of FEIs, Grant and Bauer (2004: 49-50) reject the inclusion of figurative language in the category of idioms, asserting that figuratives can be interpreted via context by recognizing the untruth and making a reinterpretation. Secondly, idioms are institutionalized phrases, that is, they should be recognized and accepted as lexical items within a certain language community at a certain time (Moon 1998: 7; Grant & Bauer 2004). In this study, inclusion in an established idiom dictionary along with frequency will be taken as evidence of the idiom being institutionalized.

In this study, idiom is defined as a semi-transparent or opaque, somewhat fixed multiword unit whose meaning cannot be derived from the meaning of its component words. For example, the idiom *jump the gun* (to do something before the right time), and the Swedish idiom *gå i taket* ('explode of anger') fill the above criteria of an idiom in this study.

4.2 Categorizations of idioms

Idiom categories have been developed using semantic, structural, and functional criteria (Liu 2008). The categories differ vastly in their restrictiveness versus openness. In this section, I will introduce Makkai's (1972), Grant and Bauer's (2004) categorizations of idioms, and Moon's (1998) categorization of FEIs. Lastly, I am going to list categories that will be excluded in this study, mainly based on Wood's (2019) categories of formulaic language.

Makkai (1972) distinguished between lexemic idioms and sememic idioms. First, lexemic idioms will be introduced. They are non-compositional in their meaning, and function as one lexeme (Makkai 1972 as referred to in Liu 2008). Below, Makkai's (1972; adapted from Liu 2008) major categories of lexemic idioms are presented with added English and Swedish examples:

1. phrasal verbs – verb + particle (carry through; köra med)
2. tournures – at least three lexons, often verb + at least two lexons (to blow off steam; *hålla huvudet kallt*)
3. irreversible binominals – two lexons in a fixed order, often nouns (fit as a fiddle; *bära eller brista*)
4. phrasal compounds – compound nouns and adjectives (hot potato; *sötebrödsdagar*)
5. incorporating verbs – compound verbs (eavesdrop; *understryka*)
6. pseudo-idioms – compound words or phrases in which one of the constituents is a "cranberry morph" or in which one or more lexons are "banned" (spick and span; *ditt och datt*)

Drawing from the Stratificational grammar, Makkai (1972) uses the term *lexon*, which in more general linguistic terms can be both a morpheme and a lexeme. All the idioms in his category of lexemic idioms are polylexonic, i.e. they consist of two or more lexons, and hence single words are also accepted as idioms. Lexemic idioms have been categorized according to structural properties of the constituent words, and there is considerable variation even within one category. For example, tournures can be constructed of verb + article + noun (to bite the dust), verb + irreversible binomial (to rain cats and dogs) and verb + direct object + modifiers (to build castles in the air) (Examples from Makkai 1972: 311-314). Additionally, Liu (2008) notes that pseudo-idioms overlap with irreversible binomials and phrasal compounds.

Sememic idioms, Makkai's second category, have a compositional meaning and additionally, an unpredictable pragmatic function (Makkai 1972 as cited in Liu 2008). Some of the functions listed in the subcategories include politeness, detachment or indirectness, proposals, interrogative greetings, and understatements. Today, most of these expressions that have clear social functions, such as the greeting "*How do you do?*", would rarely be categorized as idioms but rather as "conversational routines" or "speech formulas" among other terms (Wood 2019: 35). However, three categories of Makkai's sememic idioms are still often counted as idioms. These include "first base" idioms, i.e. expressions based on nationwide cultural institutions idioms, proverbial idioms, and familiar quotations. Liu (2008) raises two issues related to Makkai's "first base" idioms. Firstly, many would also fit in the category of

tournures. Secondly, they are not complete statements, hence diverging from the other sememic idioms. Additionally, as pointed out by Liu (2008) and illustrated by Boers, Eyckmans and Stengers (2007), other idioms, not only first base idioms, may also originate within a certain sphere of life. All in all, Makkai's broad and heterogenous categorization is not unproblematic, yet it does give a comprehensive view of the various mwus that have been, and still are counted as idioms for example in dictionaries.

Moon (1998) developed her own typology of FEIs. The first category, anomalous collocations, can be considered as an extension of Makkai's pseudo-idioms. Anomalous collocations, as indicated by the label, are syntagmatically, or paradigmatically aberrant (Moon 1998: 20). In other words, these collocations cannot be freely combined nor understood only based on their semantic associations. The first three subcategories of anomalous collocations are (1) ill-formed collocations, which break the English grammar rules (*by and large*), (2) cranberry collocations, which include items unique to that collocation (*to and fro*) (see also Makkai's pseudo idioms), and (3) defective collocations, which include senses that do not occur elsewhere (*at least*). The fourth subcategory, phraseological collocations, allow only limited variation (*on show/on display*), and therefore they are not fully productive. (Moon 1998: 20-21). When selecting Swedish idioms, I noticed that some of them were cranberry idioms (*i sänder* 'at a time', *ta något för givet* 'take for granted') that contained archaic Swedish words.

Moon's (1998) second category, formulae, is very similar to many of Makkai's sememic idioms. The common denominator is their pragmatic nature. Firstly, simple formulae have particular discourse functions, or are iterative or empathetic and syntagmatically fixed (*you know*). Secondly, sayings (*that's the way the cookie crumbles*) are famous quotations, catchphrases and truisms, i.e. stating obvious truths. Thirdly, proverbs are maxims that can be metaphorical (*every cloud has a silver lining*), or non-metaphorical (*enough is enough*). Fourthly, similes (*as good as gold*) are metaphorical comparisons identifiable by the use of the lexeme "as" or "like". (Moon 1998: 21-22).

Other categorizations have been made on semantic grounds. These include Moon's third category of metaphors, and Grant and Bauer's (2004) categories of figuratives, ONCEs, and core idioms. Moon (1998: 22-23) categorizes metaphors on a continuum of semantic transparency, with transparent metaphors in the one end and opaque, or pure metaphors in the other. Meanwhile transparent metaphors are decodable with the help of real-world knowledge (*alarm bells ring*), semi-transparent metaphors require specialist knowledge of the area the metaphor touches upon (*on an even keel*), and opaque metaphors are uninterpretable without

historical knowledge of the idiom's origin (*red herring*). (Moon 1998: 22-23). In conclusion, Moon's categorization of FEIs is much more specific when compared with Makkai's (1972), and it further specifies the peculiar characteristics of idioms.

Grant and Bauer (2004: 52) argue that previous typologies are too general, unable to distinguish various types of idiom categories, and lack specifying features of idioms. For these reasons, they developed a test for identifying "core idioms". Each level of the test is designed to distinguish a particular category of idioms. The first level distinguishes between compositional mwus and non-compositional mwus. Secondly, if it is possible to recognize the untruth in the mwu and reinterpret it correctly, the mwu is a figurative. Thirdly, if there is only one word that is non-literal or non-compositional, the mwu is an ONCE. Lastly, if the mwu is non-compositional but not a figurative, nor an ONCE the mwu is a core idiom. Additionally, it is worth mentioning that the following categories are neither included in core idioms: foreign phrase MWUs (*carpe diem*) and figurative language, including sarcasm, irony, similes, hyperbole, metonymy and metaphors (Grant & Bauer 2004: 50-52).

If only core idioms were used in this study, the number of idioms from which to choose from would be very limited. Whereas most frequent core idioms in the British National Corpus have been identified (Grant 2006), the concept of core idioms in Swedish is unknown. Consequnetly, Swedish core idioms would first have to be identified with the test. Applying the test to some Swedish is simple and clear. For instance, *en annan femma* (literally: 'another five', figuratively: 'another thing') is a clear ONCE. But further analysis reveals the main weakness of the test; it depends too much on subjective interpretation, a problem also related to Moon's category of metaphors (Moon 1998: 23). Even Grant and Bauer (2004: 55) admit that there may be borderline cases. This issue becomes evident when examining the idiom *hålla sig på jorden* (literally: 'keep yourself on the ground/earth', figuratively: 'be realistic'). One could easily see the metaphor of earth representing reality. However, for others this connection may be far from obvious, and as a result of this, other interpretations, such as earth representing humility may be invented (cf. 'stay grounded'). Similarly, *peka med hela handen* ('be particularly clear that something must be done immediately'). Some may recognize the association between hands and instructing, while for others this may never cross their minds. Additionally, even if the general meaning was recognized, it is unlikely that the specific meaning would be inferred. Due to this ambiguity, I will not use Grant and Bauer's (2004) categorization. I will only exclude figuratives that are transparent, such as *ett ljus i mörkret* ('a light in the dark').

I have chosen to exclude some of the subcategories of idioms that would well fit in the definition of an idiom. I have done this to narrow down the focus of the study, to reduce the impact that idiom variability may have on the use of strategies, and to further specify what is counted as an idiom in this study. Firstly, phrasal verbs were excluded even if they often have non-literal or figurative meanings (*hang out*), and therefore, can be counted as idioms (Wood 2015: 49). Nevertheless, phrasal verbs can be seen as a distinct group by their structural makeup (verb + preposition/particle/particle + preposition) (Wood 2019: 34). They also make up a large group of their own, as shown by, for instance, the differentiation between phrasal verb dictionaries and idiom dictionaries. Secondly, hyphenated and closed compounds were excluded since their status as multiword units is questionable. Thirdly, cranberry idioms were excluded since they may be more difficult for L2-learners.

What Moon (1998) categorizes as formula were also excluded in this study since they are generally compositional (Moon 1998: 21). Proverbs were excluded since they are often sentence-length (Wood 2019: 33), and thus syntactically more independent and, arguably, more complex, which may make them more difficult for learners. Proverbs also have a more uniform pragmatic function, which is to advise, warn, instruct, explain, or communicate common sense (Wood 2019: 33). Similes were left out since they belong to metaphors, and their structure "x is like y" often renders the idiom transparent (Wood 2019: 33). Conversational phrases and routine formulas were also excluded since they are limited to spoken language (Hult 2009: 219). Additionally, all other expressions that were full sentences, and hence could be used independently, were also excluded to allow guessing from sentence context.

4.3 Idiom studies

Idioms have been studied from a wide range of disciplines including corpus studies (see for instance Moon 1998), psycholinguistics (see for instance Conklin & Schmitt 2008), cognitive linguistics (see for instance studies of Frank Boers et.al.), and second language learning.

More generally, formulaic language has also been a research interest in pragmatics and formulaic sequences tend to have important pragmatic functions (Wood 2015: 94-95). Idioms also have a wide range of discoursal functions. Moon (1998) illustrates these functions in detail in a comprehensive corpus study. She discovered that idioms have, inter alia, informational (*behind bars)*, evaluative (*down to earth)*, and modalizing functions (*in the short run*) (Moon 1998: 219-227). Furthermore, idioms were used as politeness devices (Moon 1998: 260). This

can be explained with, for instance, their indirectness (Moon 1998: 264), and the shared familiar sociocultural schemas they evoke (Moon 1998: 267). Another interesting finding of corpus studies is that English idioms (Moon 1998) and Swedish idioms (Hult 2009: 220) tend to be common in journalism. A bunch of different types of Swedish multiword units labeled as idioms particularly occurred in texts about sports, politics, business, and economy but also in humoristic contexts (Hult 2009: 210). Additionally, both Swedish and English idioms can be found in, for instance, novels and horoscopes (Hult 2009: 210). In view of the fact that Finns have high proficiency in English (EF EPI 2021: 18), learning idioms for the comprehension of various authentic spoken and written texts and the production of rich and creative language seems reasonable.

Idioms are often combinations of familiar high-frequent words that have multiple senses and a wide range of uses (Enström 2013: 184). For instance, many idioms consist of high-frequent verbs, such as *take* or *have*. High-frequent nouns that indicate body parts, such as head or hand are also common (Niemi et.al. 2013). Research findings show that formulaic sequences, and idioms particularly are notoriously difficult for L2 learners. The difficulties learners may experience are well illustrated in a study by Martinez and Murphy (2011); even if Brazilian L2 learners of English comprehended many of the single frequent words in the text, when these exact same words were combined to form non-compositional, more or less opaque expressions, comprehension decreased dramatically. Furthermore, learners tended to overestimate idiom comprehension (Martinez and Murphy 2011), a finding also obtained related to metaphor comprehension (Littlemore et.al. 2011). The reason for this may lie in the deceptive transparency of the idioms, originating from the high frequency of single words. As a result, the learners may erroneously regard the idiom as familiar. This is further supported by, Kim (2016), who found that unfamiliar idioms that consisted of familiar words were significantly less recognized as unfamiliar when compared to single words. Additionally, the definitions of single words that were marked as familiar, were more acceptable than equivalent definitions of idioms. This may be a result of the complexity of defining idioms in general.

Macis and Schmitt (2017) investigated 170 Chilean adult English learners' knowledge of 30 polysemous collocations, that is figurative idioms. Data was gathered with a meaning-recall test and a questionnaire. The scores of the test revealed a 33.02% (9.91/30) mean of correct answers. Lower scores were obtained by McGavigan (2009 as cited in Milton 2009: 151-155) in a study of 100 Greek learners of English: even C2-proficiency learners living in an ESL-context knew little over 20% of the idioms. Additionally, McGavigan found that as the

proficiency level of the learners increased, their idiom knowledge improved (Milton 2009: 153).

In the Finnish context, studies on idioms are sparser. One of the few studies is Mäntylä's (2004) doctoral dissertation, in which she investigated how characteristics of idioms would affect their recognition and interpretation by native and nonnative speakers of English. The participants were 144 Finnish and 36 English university students. Data was gathered with a multiple-choice questionnaire. The results corroborate the findings of other studies; idioms are ambiguous, and they pose learning difficulties even for advanced L2 learners.

Native speakers favored some dictionary definitions over others (Mäntylä 2004: 99-113). This illustrates that dictionaries vary in their idiom definitions, and that some definitions may more precisely reflect the meaning of an idiom. In contrast, the definition ratings of nonnatives, were more haphazard. As for rating the suitability of idioms in different contexts, native speakers were more aware of the sociolinguistic constraints of use than nonnatives. With regards to nonnative speakers, the effects of idiom transparency, frequency and flexibility were minor. On the contrary, L1-L2 idiom resemblance was a significant factor, hence that idioms which had a Finnish translation equivalent were easiest to interpret correctly and idioms with no translation equivalents caused most difficulties. This was the case even when a semantically similar but formally different idiom existed in Finnish (Mäntylä 2004: 172).

From the data, Mäntylä induced strategies that the Finnish participants seemed to resort to when interpreting English idioms. The most prominent strategy was turning to L1 for help. The use of this strategy resulted in both positive and negative transfer (Mäntylä 2004: 173). Notably, false friends, i.e. English idioms that look similar but have a different meaning, and idioms that share one word with a different Finnish idiom, commonly led to misinterpretations. Another negative strategy that the participants would use was interpreting the idioms literally. In sum, the dissertation demonstrates that even advanced Finnish EFL-learners seem to have little overall knowledge of idioms and struggle with using transparency (Mäntylä 2004: 179) and figurative interpretation strategies as tools for interpreting the meanings of idioms. This leads Mäntylä (2004: 180-181) to conclude that the goal of L2 learners should be recognition and comprehension of idioms rather than production. This supports the choice of a receptive task for this study.

Many scholars have examined the effect different factors may have on idiom comprehension. L1 idioms and L2 idioms differ in their similarity. Their similarity can relate to either form, meaning or both aspects simultaneously. Firstly, idioms can be similar in form and meaning (*in the firing line - tulilinjalla)*. Secondly, idioms can slightly differ in form and/or

meaning (*ligga lågt* 'lie low'– *pitää matalaa profiilia* 'keep a low profile'). Thirdly, L2 idioms can be language specific, that is, lack an equivalent in L1 (e.g. *red herring, fin i kanten*). Liontas (1999; 2002 as cited in Liontas 2015) labeled the first type of idioms lexical-level idioms (LL), the second type semi-lexical-level idioms (SLL), and the third type post-lexical-level idioms (PLL). It has been established in several studies that L2 idioms with an L1 idiom identical in form and conceptual meaning (LL) are easiest to comprehend and L2 idioms that differ in form and conceptual basis from L1(PLL) are most difficult (Irujo 1986; Liontas 2002 as cited in Katsarou 2010: 123; Mäntylä 2004). Idioms that were partially similar (SLL) were also easily comprehended, but at times, L2-interference occurred (Irujo 1986; Mäntylä 2004). However, interference was not an issue with language-specific idioms (Irujo 1986).

Frequency is the most important factor that should be considered when choosing which vocabulary to learn since the higher the frequency of a word/mwu, the more often it will be encountered, and hence, the more value it has for a learner (Nation & Webb 2017). In idiom research, the effect of frequency in relation to idiom knowledge has been investigated. The findings are mixed; frequency did not have an effect, or only had a minor effect on idiom knowledge in the studies of Karlsson (2012), Macis & Schmitt (2017), and Mäntylä (2004), whereas McGavigan (2009: as cited in Milton 2009:154-155) found that idiom knowledge increased along with idiom frequency.

Steinel, Hulstijn and Steinel (2007: 456) define transparency as "the degree of overlap between the literal and the figurative meaning of an idiom". Whereas the overlap is high in transparent idioms, in non-transparent idioms the overlap is low. Thus, the figurative meaning of transparent idioms can be worked out with the help of the literal meaning since the link is obvious. The transparency of the idiom is strongly related to the compositionality of the idiom. Cieslicka (2015: 213) distinguishes between the two since an idiom can be transparent but decomposable (jump the gun) or opaque and decomposable (pop the question), and therefore, views transparency of the idiom as the idiom having a clear metaphorical motivation. However, even if in theory the decomposition of *jump the gun* from its literal meaning would be possible, it may still be highly unlikely. In several studies, transparency did not influence idiom knowledge (Karlsson 2012; Macis & Schmitt 2017; Mäntylä 2004). In contrast, transparency, as rated by university students, was found to facilitate idiom comprehension (Steinel et.al. 2007)

5. VOCABULARY LEARNING STRATEGIES

Oxford, a pioneer in the field of language learning strategies, defines language learning strategies as "operations employed by the learner to aid the acquisition, storage, retrieval and use of information" (Oxford 1990: 8). They are used to develop general communicative competence or some particular aspects of language competence (Oxford 1990:8-9). Vocabulary learning strategies (VLS) are a subcategory of language learning strategies and, therefore, only one, but a prominent branch of the language learning strategy research. However, it has to be noted that there is a great overlap between what are considered general language learning strategies and what count as VLS. The overlap can be explained by the nature of vocabulary; it is a crucial and inseparable component of language proficiency and its different aspects (Qian & Lin 2019). In the next three sections, definitions, categorizations, and studies of vocabulary learning strategies will be presented.

5.1 Definitions of vocabulary learning strategies

Strategic vocabulary learning, according to Gu (2019: 271), is an "intentional, dynamic and iterative process for the effective, efficient, and even enjoyable learning of vocabulary". Gu sees the use of VLS as part of intentional learning. Nevertheless, it can be argued that some strategies support meaning-focused implicit learning better, meanwhile other strategies are clearly more part of language-focused intentional or deliberate learning. From a different point of view, strategic actions taken by the learner may be seen as adding an intentional element to otherwise implicit learning. Furthermore, Gu (2012: 336-337, as cited in Gu 2019: 282) describes a strategic language learner as someone who actively orchestrates vocabulary learning by analyzing situated factors (context, task, learner), making choices, planning, monitoring progress, and evaluating results. The way in which Gu (2019) characterizes VLS is not as static categories, but as dynamic and flexible actions, a perspective that may easily be lost when discussing categorizations of strategies. A more pedagogically oriented definition is offered by Nation (2013: 326). In his definition, in order to deserve attention from a teacher, a strategy would need to:

1. involve choice, that is, there are several strategies to choose from and one choice could be not to use the strategy;
2. be complex, that is, there are several steps to learn;
3. require knowledge and benefit from training;
4. increase the efficiency and effectiveness of vocabulary learning and vocabulary use.

As pointed out by Gu, Nation (2013: 326) brings up learner autonomy in making choices about strategy use in the first statement. If the goal is to understand the meaning of a new word, the learner can choose between using or not using several different strategies for instance, word part analysis, dictionary or guessing from context or a combination of these strategies. The second statement implies that also within one strategy there are several steps. For example, guessing from context can proceed from the use of various contextual cues to multiple guesses and finally, confirming a guess. Nations's third statement highlights the importance of training and the role of a teacher, a view also echoed in Nation and Webb (2017), and confirmed by research findings (e.g. Mizumoto & Takeuchi 2009). Similarly to Gu (2019: 271), Nation asserts that the purpose of VLS is to increase the efficiency and effectiveness of vocabulary learning. However, it has been established that the benefits of VLS depend on the type of strategy, and the quality and quantity of the strategies used.

5.2 Categorizations of vocabulary learning strategies

In Nation's (2013: 327-330) taxonomy, VLS are divided into four different categories depending on the stage of learning: (1) planning, (2) sources, (3) processes, and (4) skill in use. Only the first and second stages are of relevance in this study. In the (1) planning stage, learner chooses what to focus on, how and when (Nation 2013: 328-329). This includes choosing words, for example based on frequency, and word aspects to focus on, strategies, and planning repetition. In the (2) sources stage, different kinds of sources, such as word analysis, context, a reference work, or using parallels with other languages, are used to find information about the words (Nation 2013: 328, 330). Typical to Nation, his categorization depicts a continuum of strategies in a multi-step process of learning a word. Merely engaging in these different stages of vocabulary learning may be strategic as such beginning not with the use of certain techniques or tactics, but with choices.

One of the most cited and widely recognized classifications of VLS is Schmitt's taxonomy (1997). It draws on Oxford's (1990) earlier classification of language learning strategies, several other sources on vocabulary learning, and survey data gathered from language learners and teachers (Schmitt 1997: 204-205). The participants were 600 Japanese EFL learners. In addition to using Oxford's (1990) taxonomy, Schmitt additionally divided the VLS into two separate stages: the discovery of a new word's meaning and consolidating a word once it has been encountered. Since the foci of this study was narrowed down to discovery strategies, only they will be presented. Discovery strategies are further divided into determination strategies and social strategies. Table 2 below presents Schmitt's VLS taxonomy for meaning discovery:

Table 2. A taxonomy of vocabulary learning strategies. Adopted from Schmitt (1997:207).

Strategy group	Strategy	Use %	Helpful %
DET	Analyze part of speech	32	75
DET	Analyze affixes and roots	15	69
DET	Check for L1 cognate	11	40
DET	Analyze for any pictures or gestures	47	84
DET	Guess from textual context	74	73
DET	Bilingual dictionary	85	95
DET	Monolingual dictionary	35	77
DET	Word lists	-	-
DET	Flash cards	-	-
SOC	Ask teacher for an L1 translation	45	61
SOC	Ask teacher for paraphrase or synonym of the new word	42	86
SOC	Ask teacher for a sentence including the new word	24	78
SOC	Ask classmates for meaning	73	65
SOC	Discover new meaning through group	35	65

	work activity		

Determination strategies are strategies used to discover the meaning of an unknown word with the help of intralingual, crosslinguistic, or contextual clues or the use of reference materials (Schmitt 1997: 208). They can be used individually, and do not involve social interaction unlike social strategies (Schmitt 1997: 205; 210). The availability of these clues is relative to the word in question, the source/context in which it occurs, and the resources at hand. With the help of grammatical knowledge, the learner may be able to designate the word into a particular grammatical category, such as verb or noun. Additionally, the learner may detect familiar affixes and roots that help to uncover the meaning of the word. Guessing the meaning may be enabled with accompanying pictures, gestures, or textual context. If the word is a cognate, i.e. word with a common parent word, it can be recognized with the help of L1, or other known languages (Schmitt 1997: 209). The same applies to loanwords. Reference materials, such as bilingual (L1⇔ L2) and monolingual (L2) dictionaries, are also helpful in meaning discovery. Since digital technologies have evolved a lot since Schmitt's taxonomy, in the present day some additions to the list can be made. In one study, where the taxonomy was used, Laffey (2020: 85) added for instance the following strategies: using a smartphone app, using an online dictionary, and using an online translator.

When searching for new vocabulary to learn, learners may use ready-made sets of flashcards (e.g. *Quizlet*) or word lists that are often developed based on word frequency, theme, function or target group. However, learners may also devise their own flashcards or lists of words based on the words they encounter, consider important, and wish to learn. Thus, word lists and flashcards may be a source of new words, a way of recording the meanings of new words, and helpful when revising the words. An equivalent category, often included in VLS surveys, is taking notes (see e.g. Gu & Johnson 1996). All in all, meaning can be discovered in various ways with the use of one or multiple strategies.

Social strategies involve consulting someone more knowledgeable, such as a teacher or a classmate. As can be seen from table 2, many of Schmitt's social strategies are restricted to the formal context of a classroom. Consulting family members or friends could as well be added to this group of strategies. In this study, the use of social strategies is not completely ruled out, but the research setting does not encourage it. Therefore, it can be hypothesized that determination strategies will predominate. Compared to Nation's taxonomy, the categories differ particularly in that there are no strategies belonging to the planning stage of vocabulary

learning. Moreover, Schmitt has clearly aimed at compiling a comprehensive list of as many types of strategies as possible.

Cohen and Wang (2018) criticize the several attempts to label and categorize language learning strategies and their functions. They argue that strategies are more complex and fluid in nature and, therefore, can fulfill several functions simultaneously (Cohen & Wang 2018: 170). Cohen and Wang's study involving six Chinese learners demonstrates that the functions (metacognitive, cognitive, social, affective) may fluctuate both within the use of one strategy, and when switching between different strategies. Strategies were also often used in sequences, pairs, and clusters rather than in isolation (Cohen & Wang 2018: 179). Thus, it seems that strategic learners can flexibly use and change the use of clusters of strategies until the task has been completed (Cohen & Wang 2018: 272). In sum, strategic vocabulary learning is a dynamic and complex problem-solving process.

5.3 Studies of vocabulary learning strategies

Studies of vocabulary learning strategies (VLS) have taken either a particular focus on a certain strategy or a more holistic focus on VLS. In this section, I will focus on studies conducted from the latter perspective. Since these general VLS studies have covered a multitude of strategies, I will particularly single out results that concern VLS related to meaning discovery. A large number of studies have aimed at investigating which vocabulary learning strategies are used and how they relate to vocabulary size. Most of the studies have centered on the VLS of EFL learners in an Asian context. These large-scale studies have revealed a positive correlation between various VLS and vocabulary size and/or language proficiency (Fan 2003; Gu & Johnson 1996; Gu 2010; Mizumoto 2010). Only some strategies, such as visual repetition, have been found to be a negative predictors of vocabulary size and general proficiency (Fan 2003; Gu & Johnson 1996; Gu 2010; Mizumoto 2010). It is also now well established from these VLS studies that dictionary use and/or guessing from context are among the most often used strategies and often considered useful (Fan 2000; 2003; Gu & Johnson 1996; Gu 2010; Kulikova 2015; Laffey 2020; Mizumoto 2010; Schmitt 1997).

Before introducing some of the VLS studies more closely, one issue that relates to the selection of VLS is addressed. Although some studies have used existing taxonomies or replicated questionnaires used in other studies (e.g. Mizumoto 2010), most studies have devised their own questionnaires (e.g. Pavicic-Takac 2008). This has resulted in various interpretations

of what counts as VLS. For example, in some studies, self-initiation and time spent on vocabulary learning have been counted as VLS (e.g. Gu & Johnson; Kojic-Sabo & Lightbown 1999). Consequently, the results have revealed general vocabulary learning tendencies and factors that relate to high language proficiency, rather than specific VLS.

In a quantitative study, Gu and Johnson (1996) surveyed the VLS of 850 Chinese second-year university students with non-English majors. The results show that, in accord with Schmitt (1997), the participants used several different guessing and dictionary strategies. In guessing from context, participants used both local and wider clues. Dictionary was used for both comprehension and learning and it entailed several look-up strategies. Participants also used notetaking as a strategy. All of these strategies, except dictionary use for comprehension, correlated positively with two dependent variables, the results of the university-wide College English Test and vocabulary size (Gu & Johnson 1996: 655). Mizumoto (2010), conducted a close replication of Gu and Johnson's study with 139 Japanese EFL students. The findings mostly correspond VLS reported by Gu and Johnson (1996). However, "meaning oriented note-taking strategies" and "using word structure" were used to a lesser extent (Mizumoto 2010: 68).

Common features of good language learners and their use of VLS have been identified in VLS research. Firstly, several studies have linked high proficiency with a conscious frequent use of a wide array of strategies (Ahmed 1989, as cited in Gu 2019: 277; Fan 2003; Gu 2003a; Gu & Johnson 1996; Kojic-Sabo & Lighbown 1999; Lawson and Hogben 1996; Sanaoui 1995). For instance, in Gu and Johnson's study, one group of learners, labeled as active strategy users, were hardworking and employed almost all strategies (Gu & Johnson 1996: 664). However, even when the range and quantity of VLS was limited, some learners were able to reach high levels (Gu & Johnson 1996; Kojic-Sabo & Lightbown 1999). Thus, although it can be generally stated that the more strategies are used the better, this does not apply to all learners, which proves that successfulness of strategies also depend on the person. This can be seen from Gu's (2003a) study of two good learners. They differed in that they treated different words with different strategies while portraying similarities in being selective of which words to learn, emphasizing multiword units, and skillful use of guessing from context, dictionary, and note-taking strategies. Moreover, data from Lawson and Hogben's (1996) study suggests that consistency in strategy use, rather than a fixed use of certain strategies is another characteristic of good language learners.

In some studies learners have completed a vocabulary learning task and additionally, think-aloud protocols or interview data have been gathered (Barcroft 2009; Gu 2003a;

Labontee 2019; Lawson and Hogben 1996). Lawson and Hogben (1996) examined the VLS of 15 learners of Italian, while they engaged in a vocabulary learning task. The data consisted of self-reports and interviews. They classified the results into four categories: (1) repetition, (2) word feature analysis, (3) simple elaboration, and (4) complex elaboration. Repetition was used exceedingly most, nevertheless, repetition strategies did not involve any significant degree of transformation (Lawson & Hogben 1996: 120). Second most, the students used simple elaboration, such as translation or simple use of context, and to a lesser degree complex elaboration, such as complex use of context or paraphrase. A correlation analysis revealed that the use of repetition and elaborative strategies were associated with recall. Yet Lawson and Hogben (1996: 121) conclude that "relatively little activity was concerned with detailed analysis of the word and its meaning in ways that would allow for the establishment of powerful associative relationships between the two". Perhaps the learners were unmotivated, unaware of deeper strategies, or simply thought that they had achieved what was required in the task.

One of the few studies that have looked at the VLS of Swedish learners has been conducted by LaBontee (2019). He gathered various data on VLS of beginner-level adult L2 learners of Swedish with various linguistic backgrounds. The analysis of interviews, surveys, and a think-aloud task revealed the use of 41 different VLS, out of which VLS for meaning discovery were most common. One interpretation of this finding could be that VLS were more focused on comprehension than deliberate vocabulary learning, which would be reasonable since the participants resided in the L2-context, rich in L2-input. Some of the highest-scoring strategies were dictionary look-up, guessing based on previous knowledge, and utilizing lexical knowledge of Swedish or other languages. Learners also tended to rely on more than one strategy. Moreover, learners who had higher Swedish proficiency and had resided longer in Sweden as well as received more Swedish language instruction had higher scores in almost all VLS categories. This indicates that the use of VLS develops along with proficiency.

Gu (2003b, 2019) argues that the choice of VLS depends on the interrelated factors related to task, learner, and context. He states that task-related factors, such as task type, materials, goal, and task difficulty affect the choice of strategies. Firstly, the task may represent intentional or implicit learning (Gu 2003b). Secondly, the learner may choose to focus on a particular dimension of vocabulary, such as breadth, depth, automaticity, or appropriateness (Gu 2019: 272). For instance, learning word lists aims at increasing vocabulary breadth, whereas writing own sentences aims at practicing the appropriate use of the word. Thirdly, learning a new word also involves several steps, such as discovering the meaning, revising the

word, or practicing its use (Gu 2003b). All these task-related factors will affect which strategies are useful and fit the purposes of the task.

Learner-related factors will also affect the way the learners approach the task since strategies are learner-initiated actions (Gu 2003b). Age-related differences in VLS have been noted in two studies (Schmitt 1997: 223-224; Gu et.al. 2005). Research also indicates that there are some sex-related differences in the number and range of VLS used (Gu 2002; Jiménez Catalán 2003; Kovanen 2014). Moreover, in one study, a moderate correlation between intrinsic motivation and VLS was found (Mizumoto 2010: 94, 133). In addition to these learner-related factors, learner styles and approaches also differ (Gu 2003a). Parry (1997) monitored the strategy use of two ESL learners in a comprehension task and found that the learners had differing approaches, holistic and analytic, both of which had their pros and cons. Similarly, Sanaoui (1995) identified a structured and an unstructured approach to vocabulary learning. In sum, learners use different VLS depending on age, sex, learning style and motivation.

The context, i.e. the learning environment, places constraints on how the task is approached (Gu 2003b). The significance of context is illustrated well in a comparative study conducted by Kojic-Sabo and Lightbown (1999). They compared the VLS of ESL learners living in Canada and EFL learners living in Northern Yugoslavia. The results exhibited some significant differences, but also striking similarities were found in note-taking strategies. The learning environment consists of people, learning traditions, curriculum, and opportunities for input and output. The amount and frequency of VLS may also change along with changes in training and the instructional setting as in the case of 100 Chinese students attending a six-month intensive English language training program (Gu 2010). Both studies point to the importance of contextual factors.

Do the VLS that are selected by learners depend on the language which is learned? Pavicic-Takac (2008: 134-145) attempted to answer this question in a cross-linguistic study of the VLS of 675 Croatian elementary students. One half of the learners studied German, and the other half English. Frequency of VLS use was measured with a questionnaire. The results show that some of the commonly used VLS, such as translation, were the same regardless of the language. Strategies that were frequently used in learning both languages could be considered as core VLS. The use of some strategies, however, varied depending on the language. This can be seen with the strategy of "listening to songs in the target language", which was much more commonly used in learning English (Pavicic-Takac 2008: 137). Due to the large sample, the study provides robust evidence that some strategies may be language-specific, whereas some

strategies are more generally used regardless of the language. Still, to what extent and how the strategy use of individuals would have fluctuated depending on the language has not yet been probed in VLS research.

The relationship between instruction and VLS is twofold. In another study, Pavicic-Takac (2008: 105-134) found that VLS were independent of the teacher's way of teaching vocabulary. Nevertheless, when the teaching has particularly aimed at teaching VLS, contrastive findings have been reported. Positive effects of explicit strategy training have been reported in some studies (Mizumoto & Takeuchi 2009; Zhao 2009), and the training regarding individual strategies such as guessing from context (Craigo et.al. 2017) and dictionary use (Ranalli 2013).

Considerably fewer studies have been conducted on the VLS with regards to idioms. One explanation to this may lie in the nature of idioms; their proportion of vocabulary is marginal when compared to single words and some other mwus such as phrasal verbs. Another explanation may be found in the common use of self-report questionnaires. Whereas learners may be able to report their general use of VLS, learners may not be so conscious of the particular VLS employed to learn mwus. Additionally, it can also be questioned whether learners put conscious effort into learning idioms in the first place. But even if there were similarities between general VLS and idiom specific VLS, and hence some research findings would be transferable, Cooper (1999) shows that idiom learning strategies also have their own unique features that relate to their specific nature. More studies on idiom specific strategies will be presented in section 6.4.

The first comprehensive VLS study that focused on idioms was conducted by Cooper (1999) who examined the processing idiom meaning. The participants were 18 nonnative speakers of English from varying linguistic backgrounds, of varying ages and with an average of 5.1 years living in U.S. Their task was to interpret the meaning of 20 idioms presented in a written context while thinking aloud. The oral answers were then transcribed and scored. Firstly, the results revealed a considerable degree of variability among the average scores of individual idioms (Cooper 1999: 244-245). In other words, some idioms were clearly easier to comprehend than others, a finding that may be explained with factors, such as transparency or familiarity with the idiom. Secondly, altogether eight strategy categories were identified. The three most used strategies were guessing from context (28%), discussion and analysis (24%), and using literal meaning (19%). Cooper's strategy of *discussion and analysis* stood for thinking aloud about the idiom or the context. Other less used strategies included: requesting information (7%), repeating, or paraphrasing the idiom (7%), using background knowledge of

the idiom (7%), using L1 (5%), and other strategies (2%). Requesting information involved asking the researcher to provide the meaning of unfamiliar single words. Using background information entailed recollecting any previous knowledge or associations with the idiom. The strategies that most often lead to correct answers were guessing from context (57%), using literal meaning (22%), background knowledge (12%), and using L1 (8%). However, only the strategies that directly lead to correct answers were counted. Thus, preparatory strategies, such as repetition, requesting information, and discussing and analyzing which likely also contributed to the correct answer, were not included in the count.

Thirdly, Cooper tested the applicability of four theoretical models of L1 idiom comprehension to the data. The comparisons led Cooper (1999) to establish that none of the models adequately explained L2 idiom comprehension, since instead of approaching interpretation linearly with a single strategy, the participants dynamically employed a variety of strategies. Cooper labeled this approach as heuristic since interpreting the idioms was tackled as a problem-solving task which involved experimenting, evaluating solutions, and trial-and-error (Cooper 1999: 254-255).

There are some issues with Cooper's study that need to be addressed. First and foremost, pre-knowledge of idioms was not checked, and therefore, existing knowledge of the idioms may have influenced the results (see also Katsarou 2010: 141-142). Instead, previous knowledge was counted as one of the strategies. There were also other methodological problems with Cooper's study. Before the actual task, the participants were given a list of "things they could talk about", that is, an account of the different strategies they could use (Cooper 1999: 241). This may have directed the participants to the use of these particular strategies, some of which may not have belonged to their regular strategy repertoire. Hence, the results may not represent the participants' actual use of VLS. Moreover, TA protocols were used but the transcriptions revealed that the researcher clearly had an active role in the TA by asking specifying questions (see e.g. Cooper 1999: 249). Research on the TA method has established that asking questions can interfere with both the participants' train of thoughts, and their task performance (more about the think aloud method in section 7.2.2

6. STRATEGIES FOR MEANING DISCOVERY

In the following three sections, I am going to examine two discovery strategies, dictionary use (6.1) and guessing from context (6.2) in more detail and present a study that focused on both of these strategies and idiom comprehension (6.3). The reason for a more thorough explanation of these two strategies is that they are among the most often used strategies by learners (Gu & Johnson 1996; Schmitt 1997: 207; Gu 2003; Laffey 2020: 87). They are often seen as complementary strategies that support vocabulary learning in different ways. Lastly, I will introduce idiom-specific strategies in section 6.4.

6.1 Dictionary use

Dictionaries can be used for comprehension in order to look up unknown words while listening, reading or translating, confirming the meanings of partly known words or confirming guesses from context (Nation 2013: 414). Research on dictionaries has focused on two different perspectives. The first perspective is lexicographical; the focus is on the dictionary producer and user. The second perspective is on language learners and learning. Both perspectives will be addressed in this section since their relationship is bilateral. For example, the status mwus have in the dictionary will affect the accessibility of dictionary information for the language learner.

Dictionaries often provide information on several different aspects of a word. These include spelling, pronunciation, definition, translation, word root, part of speech, derivations and inflections, definitions of different senses of the word, synonyms, collocations, example sentences and frequency (Laufer & Hill 2000; Webb & Nation 2017: 171). A dictionary entry for idioms may look somewhat different; some aspects may be the same as for single words whereas other aspects, such as derivations, may be irrelevant. Furthermore, dictionary entries for idioms may have distinctive features, such as etymology, images, and accounts on idiom variations. The usefulness of the different types of lexicographic information provided by dictionaries depends on the linguistic needs at the time of look-up (Heid & Zimmermann 2012: 662). The results of a couple studies suggest that meaning of a word is among the information that is commonly looked-up in a dictionary (Chan 2012: 7; Lew 2004; Lorentzen & Theilgaard

2012: 659). Additionally, Laufer and Hill (2000) found evidence that the use of multiple dictionary information supported the retention of target words. Thus, dictionary look-up may also serve as an elaboration technique that supports memorization.

Dictionaries can be monolingual, bilingual or bilingualized (Nation 2013: 423). Firstly, monolingual dictionaries are written in one language and, unless designed for language learners, understanding the extensive information provided by the dictionary requires good L2 comprehension skills. Secondly, bilingual dictionaries, are written in two languages, which make them more L2-learner friendly. They can be bidirectional (L1→L2, L2→L1). Thirdly, bilingualized dictionaries are a combination of the two, including both a translation of the headword and its senses, and the information provided by a monolingual dictionary (Lew & Adamska-Salaciak 2015: Nation 2013: 423-425). Moreover, there are learners' dictionaries, which do not contain low-frequency vocabulary (Nation & Webb 2011: 60). The types of dictionaries have been assessed and compared in several studies. The findings suggest that learners prefer bilingual dictionaries (Atkins & Varantola 1998a: 43; Lew 2004), which also support comprehension best (Lew 2004), and are most useful and most used in looking up unfamiliar vocabulary while translating (Atkins & Varantola 1998b: 99, 111). Moreover, the findings of Laufer and Melamed (1994) and Laufer and Hadar (1997) imply that overall dictionary skills may determine which dictionary is most useful, but for all learners regardless of their dictionary skills, bilingualised dictionary proved to be superior for the comprehension of words. The predominance and superiority of bilingual/ized dictionaries are evidence of the fundamental role of L1 in foreign language learning. However, the use of monolingual dictionaries seems to increase along with L2 proficiency (Atkins & Varantola 1998a: 43; 1998b: 113; Lew 2004). Likewise, learners with better dictionary skills, may be able to make better use of monolingual dictionaries when comprehending unfamiliar words (Laufer & Hadar 1997; Laufer & Melamed 1994).

Today, electronic dictionaries (e-dictionaries) or web-based online dictionaries that can be accessed via different kinds of devices, have become more and more common replacing or complementing printed dictionaries (e.g. Kosem et.al. 2018; Lew & de Schryver 2014). There are several benefits with e-dictionaries. Firstly, they offer easy and fast access to information, which may encourage multiple searches of information.Secondly, multimodal elements such as audio, visuals, video or mini-concordances; i.e. text lines taken from a corpus, can complement a word entry (Nation 2013: 425). E-dictionaries can also be accessed and found via several ways. These routes include entering a query into a search engine and clicking on the link appearing in the result page, typing the site URL or clicking on a referring link on

another website (Lorentzen & Theilgaard 2012: 654). Today, traditional and conventional English dictionaries, such as *the Cambridge Dictionary*, *Oxford Learner's dictionary* and *Macmillan Dictionary* are freely available in online. Similarly, the combination of three Swedish dictionaries, *Svenska Akademiens Ordböcker*, and the high-quality bilingual dictionary *Lexin* is available for e-use (for more about Lexin see Hult et.al. 2010).

Learning-oriented studies on dictionary use, both with a focus on single words and collocations, abound. Many studies on single words have shown that dictionary use can lead to learning single words (Knight 1994; Hulstijn et.al. 1996; Li 2010; Luppescu & Day 1993; Mekheimer 2018). Studies focusing on mwus have, for the most part, focused on collocations and more specifically, the written production of missing collocates with the help of a dictionary. It has been established that access to dictionaries improves written production of collocations (Chen 2016; Dziemianko 2010; Heid & Zimmermann 2012; Laufer 2011). However, in these studies, dictionary use did not guarantee success in collocation production, since availability and location of dictionary information varied and the learners had insufficient dictionary skills, both of which may lead to a failure in dictionary search.

One of the few studies on meaning determination of mwus has been conducted by Li and Xu (2015). They investigated 32 Chinese EFL students' use of the Macmillan English Dictionary Online in determining the meaning of polysemous verb phrases. The data consisted of pre- and post- meaning determination tests and intro-and retrospective questionnaires. The results show that dictionary consultation improved the results on the meaning-determination test and that especially the dictionary definitions and examples were regarded as useful by the participants. However, dictionary consultation was not unproblematic. The problems identified by the participants related to either the information provided by the dictionary, the participants' knowledge of words, or the participants' dictionary skills. Although the study has several similarities with this study, there are also several differences such as the restriction to use only one monolingual dictionary and the selection of polysemous verb phrases as the target phrases. Nevertheless, idioms that can be interpreted both literally and figuratively could also be considered polysemous.

Successful dictionary use requires dictionary skills. In 1982, Scholfield argued that dictionary use is not a straightforward activity since it involves several steps, requires prior knowledge of, for example, English language and dictionary conventions, involves a number of stages in which hypotheses are set up and tested, and is tied to lexical and contextual inferencing (Scholfield 1982: 185). With the change from the use of print dictionaries to online and electronic versions, some of these skills, such as understanding alphabetization, have

become irrelevant, whereas new skills, specific to online dictionaries are needed (Lew 2013: 19). This shift in dictionary skills is demonstrated by Lew (2013), in his overview of Nesi's (1999) exhaustive list of dictionary skills for higher education. Three of the altogether six upper stages of Nesi's (1999) list are: 1) before dictionary consultation, 2) locating entry information, and 3) interpreting entry information. Stage two includes the assessment of the necessity of dictionary use, i.e. cost vs. benefits, deciding what to look up, in what form, and in which dictionary. Contextual inferencing is also included in this stage. Stage two includes understanding dictionary structure and conventions and searching and locating skills. Stage three skills include navigating dictionary information, skills in extracting relevant information, and skills in interpretating various dictionary information and applying it. (Nesi 1999; Lew 2013: 17-22). Most of these stages overlap with the steps hypothesized by Scholfield (1982), and the strategies that five "good language learners" used in Koyama and Takeuchi's (2009: 140) study involving reading while thinking aloud and using an e-dictionary.

Several difficulties that learners face when using a dictionary have been identified and described in literature. To date, several studies have confirmed that dictionary use does not guarantee success in linguistic tasks (e.g. Atkins & Varantola 1998a: 40; Prichard & Matsumoto 2011). Firstly, learners may ignore information or struggle with finding the correct information, and not be satisfied with the information they found (Atkins & Varantola 1998b: 101; 103-104; Nation & Webb 2017: 172; Wingate 2004). Secondly, lack of vocabulary may inhibit the learners from understanding monolingual definitions or example sentences, which are often very complex (Lew & Adamska-Salaciak 2015: 50; Nation & Webb 2017: 172). Additionally, the learners should be capable of transforming monolingual definitions and integrating them into the context (Wingate 2004: 10). Thirdly, many words are polysemous which leads to difficulties in choosing the meaning that is searched for among multiple senses (Chen 2016; Lew & Adamska-Salaciak 2015: 51; Nation & Webb 2017: 172). Finally, learners may use dictionary excessively by looking up irrelevant words (Prichard 2008: 221; 226).

Wingate (2004) investigated 17 Chinese learners of German and their use of several dictionaries while reading. She concludes that learners lack basic strategies required for dictionary use and instead, use inappropriate strategies that lead to failure (Wingate 2004). One of these negative strategies was the use of a kidrule strategy which implies "selecting familiar segments from an entry and inserting them in the text" regardless of their appropriateness (Wingate 2004: 5, 9). Another negative strategy, that was also identified in a study on collocations (Chen 2016), was choosing the first entries in a list of longer entries (Wingate

2004: 5, 8). As a consequence, Wingate advocates instruction and practice of dictionary use (Wingate 2004: 11).

Some dictionaries do not support the search of mwus or have not included a specific mwu, which may lead to an unsuccessful search in an e-dictionary (Lorentzen & Theilgaard 2012: 656, 659). Another encountered problem in locating mwus is that the typed form of the multiword unit may differ from the lexical form appearing in the dictionary (Dwornik and Margol 2011, as quoted in Lew 2012: 7). These idiom variations may also entail changes in meaning, which may result in two different failures (Szczepaniak 2006: 57-58). Firstly, the dictionary definition may be rejected as inappropriate when compared with the context. Secondly, the dictionary definition may not be adjusted and applied to the context. Thus, the semantic distance between the canonical form and the idiom variation will affect reception even when dictionary definition is found (Szczepaniak 2006: 45). However, e-dictionaries can support the locating of multiword units with different kinds of search interface enhancements, such as displaying a menu of suggested mwus and showing snippets of other related dictionary entries including the target lemma (Lew 2012: 11-13).

Prestigious and conventional dictionaries, either paper or online versions, are not the only source of information on words (Yongwei 2012: 423), and when lexical problems are not solved using a dictionary, the most popular alternative is to consult other web sources (Kosem et.al. 2018: 108, 110-111). Nesi (2012) scrutinized less prestigious "alternative e-dictionaries" (AEDs), which are far less researched and discussed despite their popularity. AEDs are usually a combination of diverse dictionaries, some of which may be from highly regarded publishing houses, whereas some may have their origin in other online sources such as the collaboratively edited *Wiktionary*. Yongwei (2012: 423-325) refers to the former as one-stop dictionary sites and the latter as DIY dictionaries. Nesi's (2012) analysis shows that AEDs lack explicitness, transparency, and consistency in many aspects such as the author, guiding principles, or record of change. Moreover, as Nesi (2012) summarizes, studies of AEDs have found their quality to be unreliable. The flaws found in studies include errors, ungrammatical inflections, outdated language, taboo language and narrow entries.

In the same vein, Moon (2015) compared the information about idioms provided by non-lexicographical and non-academic websites with learner dictionary entries. Two idioms, *hit the nail on the head* and *kill two birds with one stone*, were among the 16 items appearing in more than one web source. These were selected for a more thorough analysis. The examination revealed that the information found on the chosen websites was less structured and less satisfactory in terms of explanation and exemplification when compared to

conventional establishment dictionaries (Moon 2015: 330). Nevertheless, Moon (2015) concludes that web searches may prove to be a fast pass to user-friendly and sufficient information about the idiom in question, although the definitions and examples found can at times be deficient and misleading. Moreover, Moon (2015: 330) remarks that comments and discussion on individual idioms can be found on many other websites that focus on certain aspects of the idiom. DIY-dictionaries, such as *Wiktionary* and *Urban Dictionary* also have their advantages, which include unlimited ability to add up-to-date entries and provide several examples and additional information (Yongwei 2012). In conclusion, it can be stated that language learners may find satisfying information on words online, but they should be critical when consulting non-established dictionary-like sources.

6.2 Guessing from context

Guessing from context refers to inferring the meaning of unknown or partially known words from input with the help of linguistic clues, background knowledge, or common sense (Webb & Nation 2017: 93). It is one of the most commonly used vocabulary strategies since it accounts for the majority of vocabulary learning in L1 (Nation 2013: 348). A substantial amount of research has been done on the topic since the 1970s. Guessing from context is meaning-focused and, therefore, it can be considered as a form of implicit vocabulary learning (Webb & Nation 2017: 93). However, for guessing to be successful, noticing has to occur (Webb & Nation 2017: 94). The mode of input can be listening, reading, or doing both simultaneously, for instance when watching a movie. In this section, I will focus particularly on reading, since guessing from context is also considered to be a reading strategy and consequently, it has received a lot of attention from both L1 and L2 perspectives. Most of the studies focus on English, but some other languages have also been researched.

The context can provide a wide array of different types of clues. In an early study Ames (1966) observed L1 Ph.D. students' use of contextual clues in reading. As a result of the study, a comprehensive classification scheme of contextual clues was compiled. The clues are listed below in the order of Rankin and Overholser's (1969) rankings of effectiveness (as quoted in Nation 2013: 365):

1. Words in series
2. Modifying phrases
3. Familiar expressions

4. Cause and effect
5. Association
6. Referral clues
7. Synonym clues
8. Definition or description
9. Preposition clues
10. Question and answer
11. Comparison and contrast
12. Main idea and supporting details
13. Non-restrictive clauses

In another taxonomy, Bengeleil and Paribakht (2004) identified linguistic and non-linguistic sources for guessing from context. Linguistic sources were divided into intralingual and interlingual sources. Intralingual sources were further divided into target word level, sentence level, and discourse-level clues. Interlingual sources included lexical knowledge and knowledge of word collocation.

Clarke and Nation (1980) proposed the following inductive five-step technique for guessing meaning from context: determining the part of speech, looking at the immediate grammatical context, studying the wider context, guessing the meaning, and lastly, checking the guess. Nation (2013: 377-379) instructs that looking at conjunction relationships, such as restatement, time sequence, or summary, may be helpful during studying the wider context. Analyzing word parts, i.e. roots and affixes, should be included in the fourth step. However, this is only a proposal and in reality, learners may take different routes to infer the meaning of unknown words as shown in studies that compared different approaches taken by learners (Arden-Close 1993).

Many different aspects of word knowledge can be learned from context. Most often attention is given to learning the form-meaning connection of the word. In particular when languages are culturally close, the learner already has an existing concept for the newly encountered form, however, when the concept is unfamiliar, inferring the meaning becomes more difficult (Nagy et.al. 1987). This may be the case with some Swedish and English idioms that lack semantic idiom equivalents in Finnish, such as *hot air* or *go down in flames.* In addition to the form-meaning connection, learning from context can have an enriching function on partially known words (Nation 2013: 359). Thus, the potential aspects for learning can be various and for instance include learning collocates, affixes, various senses of a word, and the

grammatical functions of the word. Likewise, idioms may be recognized as a string of words that co-occur, even if the meaning is not yet discovered.

As pointed out earlier, vocabulary learning is incremental and gradual in nature, and this is particularly true to implicit learning from context (Webb & Nation 2017: 55-56). Sometimes guessing the meaning of a word is incomplete, which is illustrated by Clarke and Nation (1980: 213-214), who propose that guessing with the help of the immediate and wider contexts may only lead to determining the general meaning or establishing an idea or connotation (e.g. positive or negative) of a word instead of its specific meaning. These increases in knowledge may contribute to eventual learning of the word, and should, therefore, not be dismissed (Nation 2013: 366). Frantzen (2003: 173) also raises an important point about distinguishing between inferring and learning. Although both are relevant components in vocabulary acquisition, a correct inference does not necessarily result in learning.

Previous research has established that L2 learners can learn aspects of meaning from context (e.g. Day et.al. 1991; Brown et.al. 2008; Pigada & Schmitt 2006; Waring & Takaki 2003), nonetheless, vocabulary gains are often fairly modest (e.g. Brown et.al. 2008). Much of the results obtained from other early studies focusing on L2 learners have also found positive effects on learning words from context. However, the results of these studies must be interpreted with caution due to inadequacies in methodology (Waring & Takaki 2003; Nation 2013: 356). The findings of Waring and Takaki's (2003) carefully implemented study indicate that new words can be learned from reading graded readers. However, learning new words by guessing from context was proportionally slow and ineffective. Additionally, learning rates decreased over time, but this may have been due to not being able to encounter the target words after reading because disguised (non-existing) word forms were used in replace of real existing concepts. The study also confirmed the much-voiced observation that successful learning of words from reading requires multiple encounters with words, i.e. large quantities have to be read to learn new words (Nation & Wang 1999).

Many different variables have been found to have an effect on how successful vocabulary learning from context is. Contextual factors include informativeness of the context, i.e. the number of contextual clues (e.g. Webb 2008b); density of unknown words (Liu & Nation 1985), and the number of encounters with the word (e.g. Brown et.al. 2008; Chen & Truscott 2010; Pigada & Schmitt 2006). Learner-related factors include vocabulary size (Shefelbine 1990), the depth of vocabulary knowledge (Nassaji 2006; Shefelbine 1990), familiarity with the topic, and knowledge of passage sight vocabulary (Pulido 2007), language proficiency (Zahar, Cobb, & Spada, 2001) and the capacity of working memory (Daneman &

Green 1986). Lexical factors include the existence of a translation equivalent in L1 (Chen & Truscott 2010) and the conceptual difficulty of the words (Nagy et.al. 1987). Other possible mediating variables, among others, include the proximity of relevant contextual clues, whether the word is concrete or abstract, amount of polysemy, and familiarity of the concept (Jenkins & Dixon 1983: 250-253).

Many studies have focused on L2 learners' approaches to guessing from context. Nation (2013: 367) summarizes the results of these studies into the following actions taken by a good guesser: (1) using a variety of clues, (2) checking various types of clues against each other, (3) not letting the form of the word play too large a part, and (4) not arriving at a guess prematurely. The third characteristic is well illustrated in Frantzen's (2003) study, where some of the students' incorrect guesses were based on formal similarities between L1 or L2 words. Moreover, the study illustrated well how inattentive reading resulting in ignoring some of the contextual clues and oblivious certainty of the meaning can lead to unsuccessful guesses. In conclusion, successful guessing from context is a careful multi-step process.

Studies on learning multiword units are sparser and only recently more attention has been directed to this area of research. Most of the studies have focused on collocations. Both transparent (Durrant & Schmitt 2010; Szudarski & Carter 2016; Pellicer-Sánchez 2017), and opaque or figurative collocations (Macis 2018; Webb et.al. 2013) have been researched. As with the studies including single words, research shows that L2 collocations can likewise be learned from context (Macis 2018; Pellicer-Sánchez 2017; Webb et.al. 2013). Most studies have focused on frequency effects. The findings are mixed; both positive effects (Durrant & Schmitt 2010; Webb et.al. 2013), arbitrary effects (Macis 2018), and no significant effects (Pellicer-Sánchez 2017) of frequency have been found. The results of Pellicer-Sánchez (2017) may be explained with the use of only 4 or 8 occurrences of the target collocations, when for instance compared with Webb et.al (2013) 5, 10, and 15 encounters with the target collocation. However, Webb et.al. (2013) used the reading-while-listening input mode which has been found to be slightly more effective for vocabulary learning than the reading-only mode (Brown et.al. 2008). Moreover, the text was a graded reader, which may have facilitated learning. Despite these findings, Szudarski & Carter (2016) found that only the condition of input flood + input enhancement led to acquisition of collocations. Concerning the quality of repetition, Durrant and Schmitt (2010) found that verbatim repetition had a greater impact on collocation learning than varied repetition. Overall, there seems to be some evidence to indicate that increasing the prominence of collocations in a text may increase the chances of learning from reading due to the low frequency of collocations in natural texts (Boers 2019: 146).

In three case studies, Macis (2018) investigated L2 learners' knowledge of figurative collocations before and after reading a long novel. The novel was semi-authentic since 38 verb-noun and adjective-noun collocations had been inserted in it. Participants were three advanced L2 learners. Based on pre-and post-interviews, the participants' answers that indicated partial or full knowledge of the figurative meaning were scored. All participants had made strong and durable vocabulary gains in terms of figurative collocations. The interview responses also revealed that participants had used the guessing from context strategy successfully, but also tried to unsuccessfully infer the figurative meaning from the literal meaning (Macis 2018: 60-61). Some of the initial wrong interpretations in the pre-test were sustained despite frequent occurrences in helpful contexts. This may be the result of an enduring memory trace (Macis 2018: 61). The pre-test may also have had some effect on the learners' awareness of these collocations. In the case of one particular collocation, *Big Brother*, Macis (2018: 61) speculates that the lack of cultural knowledge may explain the misinterpretations. All in all, the study gives relevant insights into learning idioms from context, nevertheless, it must be reminded that the positive results reflect the learning of solely three advanced L2 learners.

Several studies have acknowledged the relevance of the guessing from context strategy in interpreting the meanings of unknown idioms (Cooper 1999; Katsarou 2010; Liontas 2003; Wray et.al. 2016; Xie 2017). In two of the studies, guessing from context was found to be the most used strategy in interpreting the meanings of idioms (Cooper 1999: 246; Zuo 2008 as quoted in Wray et.al. 2016: 6). Wray, Bell, and Jones (2016) compared 17 native and 23 non-native adult English speakers' interpretations of unfamiliar formulaic sequences from Georgette Heyer's novel. Not only did the native speakers use more context, but they also used more analogies in their interpretations, something that was found to increase along with proficiency. Non-native speakers, on the other hand, made more references to unknown words which may indicate that they used a less holistic strategy. Wray et.al. (2016) speculate that this may be due to lower language proficiency, lack of confidence, or interestingly, an approach where meanings of single words are used to decode the meaning of the whole.

Probably the most comprehensive study on idioms and guessing from context has been conducted by Katsarou (2010). Her doctoral dissertation that centered around the inferencing strategies of 60 Greek EFL students. The data consisted of an idiom identification test in which unknown idioms had to be underlined, and a multiple-choice test on text comprehension. Additionally, participants filled in a mini-questionnaire which involved indicating familiarity with the given idiom, guessing its meaning, and marking which of the seven strategies on the list were used in guessing. Other data was also gathered about language proficiency, learning

style, and background information. The results of the identification task show that identifying L2 idioms as unknown lexical units is a difficult task since the percentage of identification was 13.12 across all types of idioms (Katsarou 2010: 286). This finding, together with the participants over-confidence of their comprehension (Katsarou 2010: 291) corroborates Martinez and Murphy's (2011) findings. Of multiple factors, only low frequency of the words within the idiom and low interlingual similarity helped in identifying the idioms as unfamiliar (Katsarou 2011: 286-287).

The meaning of the idioms turned out to be difficult to guess, as shown by the low 37.29% mean rate of accurate guesses (Katsarou 2010: 288-289. On a more positive note, the results could be seen as evidence of the usefulness of guessing strategy. The accuracy in comprehension correlated with interlingual similarity hence that L2-L1 identical idioms were easiest to comprehend (67.7%) followed by similar idioms (56.7%), and different idioms were the most difficult to comprehend (19.7%) (Katsarou 2010: 290-291). These findings are in line with earlier research. Moreover, Katsarou compiled a list of "clues" or guessing strategies. Below this list is presented according to the frequency of use, which also is identical to the order of the successfulness of the strategy (Katsarou 2010: 172, 293, 297-298):

1. meaning of the sentence (57.33%)
2. meaning of the words surrounding the idiom (43.59%)
3. words in the idiom itself (28.47%)
4. idiom metaphor (22.48%)
5. meaning of the paragraph and/or the whole text (15.05%)
6. an equivalent idiom in Greek (11.67%)
7. background knowledge of the topic (2.14%)

As can be seen, strategies 1, 2, 3, 5 and 6 are linguistic sources of information, whereas strategies 4 and 7 are non-linguistic (see Bengeleil & Paribakht 2004: 231). Altogether the Greek students' strategy use was low; that is, only few strategies (1.7 strategies) were employed per unknown idiom (Katsarou 2010: 293). Furthermore, the successful use of strategies was extremely low; only 11.69% of the employed strategies led to comprehension (Katsarou 2010: 297). These findings can be taken, not only as evidence of the low quality of strategies, but also as evidence of the trial-and-error nature of the strategy use, which involves trying and testing until a helpful strategy is found (see also Cooper 1999: 245-255). However, the results concerning the guessing strategies should be viewed in the light of two considerations that may

have affected the results. Firstly, the type and number of strategies may have differed in reality, since the participants may have used other strategies, for example the use of other known languages, that were not included in the list. Secondly, seeing the ready-made list of strategies may have steered the participants' actions. Still, in accord VLS research, a significant positive correlation was found between the use of strategies and successful L2 idiom comprehension, hence that the use of more and varied strategies lead to more successful guesses (Katsarou 2010: 295).

6.3 Idioms: dictionary use and guessing from context

Despite the magnitude of studies investigating dictionary use and guessing from context and even the more recent interest in the study of mwus, less studies have focused particularly on idioms. One study that fills this gap was conducted by Szczepaniak (2006), who investigated the role of a monolingual dictionary in the comprehension of modified idioms. Guessing from context was also examined from a comparative viewpoint. In her study, 75 advanced Polish EFL students, divided into two groups of which only one was allowed to use a dictionary, received five excerpts containing five low-familiarity modified idioms, i.e. non-canonical idioms (e.g. have a *technical* axe to grind*)*. After reading the excerpts, participants paraphrased the meaning of the underlined sentences that contained the target idioms, formulated idioms using word cues (words that occur both in the canonical and the modified form), and wrote a definition for the idiom. The answers were scored with the help of an idiom correctness scale. The dictionary group also filled in a self-report table on dictionary look-up behavior.

First and foremost, when the results of the five idioms were compared, considerable alteration was found. Szczepaniak (2006) reckons that the alteration may depend to a lesser degree on (1) the participants' familiarity with the idioms (Szczepaniak 2006: 70), (2) the extent to which the idiom has been modified (Szczepaniak 2006: 81), and to a higher degree, (3) contextual factors. Another significant finding was that context alone was a limited predictor of comprehension. This accords with Frantzen's (2003) findings on guessing single words. Similarly, both Frantzen's (2003) and Szczepaniak (2006: 76) observed that the context could even dissuade the learner from the correct meaning of a word. In Szczepaniak's (2006) study, comprehending the contextual meaning often failed or was imprecise. Again, several different reasons were found to account for these findings. Contextual factors include the informativeness of the context and dual activation, i.e. both literal and figurative meanings fit

in the context (Szczepaniak 2006: 77). Reasons that relate to the participants include a lack of world knowledge and negative transfer (Szczepaniak 2006: 77). Additionally, some participants resorted to negative strategies such as perfunctory reading and overstressing other comprehended meanings (Szczepaniak 2006: 78). Furthermore, the idiom formulation and definition tasks revealed that the forms and definitions of the "guessing" group were often non-canonical and entangled with the context (Szczepaniak 2006: 78-81). From this, Szczepaniak (2006: 81) concludes that the degree of contextual transformation of the idiom will affect inferring the canonical meaning and form.

The participants in the dictionary group chose to utilize a dictionary with a 78% rate (Szczepaniak 2006: 70). However, the number of consultations did not necessarily lead to improvements. The most common reason for dictionary consultation was unfamiliarity with the idiom (63%), but it was also used to confirm the meanings of familiar idioms (16%) (Szczepaniak 2006: 72). Some participants chose to disregard the opportunity for dictionary use (21%), even when idioms were unfamiliar (15%) (Szczepaniak 2006: 72). For the most part, there was no trouble in locating the right idioms in the third edition of the *Longman Dictionary of Contemporary English*. The few failed consultations could be attributed to an inability to use cross-references and the use of wrong or nonexistent headwords (Szczepaniak 2006: 81-82). The dictionary was used to both form new ideas (72%) and uphold old ideas (28%) (Szczepaniak 2006: 74). An examination of the effectiveness of dictionary use reveals that "dictionary use yielded negative results in 72% of the cases by fostering or corroborating erroneous ideas and generated or supported positive ideas in merely 28%" (Szczepaniak 2006: 74). Getting stuck to initial incorrect inferences despite dictionary look-up has also been observed by Macis (2018). From this, it can be concluded that access to a monolingual dictionary does not guarantee comprehension in the case of modified idioms.

There are several explanations as to why successful consultations did not guarantee comprehension. Firstly, with two idioms, in particular, the dictionary definitions turned out to be peculiarly phrased and inadequate in capturing the complex meanings of the idiom. For instance, in one of the definitions a certain aspect was overemphasized (Szczepaniak 2006: 82). Comprehension could also have been boosted if the entries had included illustrating examples (Szczepaniak 2006: 83). Secondly, participants used the dictionary poorly. Participants often failed to reconsider the context in order to apply and adjust the dictionary meaning. Moreover, participants tended to take only parts of the definition and dismiss the whole. Nevertheless, dictionary look-up did improve comprehension (Szczepaniak 2006: 67). The written task shows that the exposure to a dictionary definition also enhanced the participants' awareness of

the canonical form and the modified form of the idiom (Szczepaniak 2006: 84). All in all, Szczepaniak (2006) concludes that guessing from context and dictionary should be used in tandem.

It can be hypothesized that some adjustments will make idiom comprehension easier in this study when compared to that of Szczepaniak (2006). Firstly, the participants will be allowed to use many different online dictionaries, both mono- and bilingual ones. Secondly, participants will be given the canonical form of the idiom. Thirdly, the contextual sentence will be taken from the example sentences given by a dictionary, which usually employ the canonical meaning of the idiom.

6.4 Strategies for learning idioms

In addition to the strategies put forth by VLS-research, idiom-specific strategies have emerged from other studies, particularly within the branch of cognitive linguistics. These studies have centered around investigating the pedagogical value of different methods for learning idioms. In this section, I will introduce and review three strategies that can be used in meaning discovery, opting out strategies that solely focus on memorization. These are: using images/pictorial elucidation, etymological elaboration, and using metaphor. The underlying argument that is often presented for the use of these strategies is that idioms are not arbitrary but motivated, i.e. they can be explained, in various different ways (Lakoff & Johnson 1980; Kövecses & Szabo 1996; Boers et.al. 2007; Boers & Lindstromberg 2008). Researchers have also proposed the Dual Coding Theory (Paivio 1971), and the Levels of Processing Theory (e.g. Craik & Lockhart 1972) as a theoretical rationale for these strategies. In a nutshell, these theories rest on premises according to which associating verbal information with mental images, and deep-level processing enhance memorization (Boers & Lindstromberg 2008:11-12). However, unlike many other VLS, these strategies cannot be applied to all idioms, but their utility depends on the characteristics of the idiom, for example, its figurativeness.

Pictures can be used as support for idiom learning. This strategy has also been called “pictorial elucidation” (Boers et.al. 2009) and “metaphorical elaboration” (Ramonda 2016). Idioms vary a lot in how imageable they are (Ramonda 2016: 97). Some idioms may be particularly imageable and lend themselves very well to a graphic presentation (Szcepaniak and Lew 2011: 323, 342), whereas other idioms may be hard or even impossible to illustrate. One of the underpinnings for the use of idiom pictures lies in the positive effect imageability

can have on learning idioms (Steinel et.al. 2007). Szczepaniak and Lew (2011: 331) list different possibilities for idiom illustrations:

1. both the literal and figurative layers of meaning are depicted (may include etymology)
2. only the idiomatic meaning is illustrated
3. only the literal reading is represented
4. the literal meaning is humorously exploited

This list demonstrates that the choice of what to illustrate is by no means straightforward. A quick search using Google images reveals that diverse illustrations of idioms are available online. Up to now, a several studies have demonstrated that verbal explanations of meaning accompanied by images facilitate recognition and recall of meaning (Aydin 2017; Boers et.al. 2008: 203-205; Ramonda 2016; Szczepaniak & Lew 2011). As for retention of form, Szczepaniak and Lew (2011) found a positive effect, whereas no such effect was found in the studies of Boers and colleagues (Boers et.al. 2008: 203-205; Boers et.al. 2009), which raises speculations about the possible distracting role of pictures on the learning the form of the idiom.

In most studies that have looked at the effect of images on learning idioms, literal meanings of the idioms have been depicted. The effect of the type of picture and idiom transparency on meaning recall was investigated by Ramonda (2016: 142) in one of his many studies. In his study, 64 Japanese university students took part in learning sessions where they learned the meanings of 27 idioms in one of the following conditions: no image, image of the literal meaning, image of literal + figurative meaning. After exposure to one of these conditions, they completed immediate and delayed recall tests. It turns out that both types of images had a positive effect on meaning recall. Ramonda (2016: 123) postulates that even literal pictures may draw more attention and encourage deep processing. Also, the literal + figurative condition was superior to only literal condition, particularly for high transparency idioms. Instead, for mid-and low-transparency idioms, scores were fairly similar. Interestingly, the scores of individual idioms varied a lot, which leads Ramonda (2016) to suspect that the learning burden of idioms is affected by many other factors than the presence of images, transparency, and L1-L2 equivalence. It was also established that occasionally both types of pictures were misleading, for instance when leading to an over-specification of the meaning.

Boers and colleagues (2004; 2007) have introduced etymological elaboration as a mnemonic strategy for learning idioms. It involves raising awareness of the literal origins or the source domain of the idiom. Some common source domains of idioms include

fighting/warfare (e.g. *in the frontline)*, food/cooking (e.g. *bread and butter*), and games/sports (e.g. *hedge your bets*) (Boers et.al. 2004: 59). The etymology of an idiom can also entail historical and cultural knowledge (e.g. *red herring*). Etymological elaboration is an instance of semantic elaboration, which according to Cohen et.al. (1986 in Boers et.al. 2007: 45) involves "active and rich processing of an item with regard to its meaning". Boers et.al. (2007: 56) present a 4-step method for learning idioms. If the assistance of a teacher is ruled out, the steps would be the following: (1) hypothesize about the origin of the expression, (2) refine or rectify your hypothesis, (3) interpret the meaning by combining etymology and context, (4) refine or rectify your interpretation.

Boers et.al. (2007) conducted an experiment with two groups of Dutch college students majoring in English. Both groups completed on-line exercises including a multiple-choice meaning identification task. However, prior to this task, the experimental group identified the source domain of the idiom and thereafter received feedback in the form of a short etymological note. The results show that etymological elaboration facilitated meaning recognition. Other studies involving Iranian EFL-learners have confirmed the positive effects of etymological elaboration for the learning of idioms (Bagheri & Fazel 2010; Haghshenas & Hashemian 2016; Noroozi & Salehi 2013). But in contrast, Szczepaniak and Lew (2011) did not find any positive effects of etymological notes. However, the results, indicating that the presence of an etymological note was distracting, can be explained with the differences in the nature of the activities used in the study; in Boers et.al. (2007) study the participants hypothesized about the origin of the idiom which likely contributed to a higher cognitive effort when compared with reading etymological notes independently (Szczepaniak & Lew 2011: 342). Moreover, Vasiljevic (2015) compared the effect images and etymological notes had on receptive and productive retention of idioms. She found that pictures of the literal meaning benefited more the production of idioms, whereas etymology led to better retention of the figurative meaning. From this, it can be hypothesized that literal images may draw more attention to the form of the idiom.

Most of the mentioned studies have focused on retention of meaning and/or form. A welcome exception to this is one of Karlsson's (2019) studies in which she compared the effects of context, pictures, etymological notes, and pictures + etymological notes on the comprehension and retention of meaning. The participants, 120 Swedish university students, were assigned to two groups. The first group was exposed to the target idiom in one of the four conditions and then had to guess the meaning of the idiom. Unlike in most reviewed studies, neither explanations nor options for meaning were given (see Boers et.al. 2007). The second

group also received treatment in one of the four conditions but additionally was exposed to the correct meaning. After this, the second group completed immediate and delayed meaning retention tests. The results show that the group who saw the idiom in context outperformed all other groups in the comprehension test, and the difference was statistically significant. On the retention test, no statistical significance was found between the context, picture, and etymology groups, however, for long-term retention, pictures and etymology turned out to be more beneficial. This is in line with the findings of other studies. With regard to both groups, the helpfulness of the clues, that is pictures that alluded to the figurative meaning and etymological notes that supported the figurative meaning proved to play an important role in comprehension and meaning retention.

Many idioms are metaphorical. In a classical paper, Lakoff and Johnson (1980) argue that there are conceptual metaphors (CMs), such as *argument is war*, *love is a journey*, and *time is a moving object* that are systematically reflected in everyday language, including idioms. According to them "the essence of metaphor is understanding and experiencing one kind of thing or experience in terms of another" (Lakoff & Johnson 1980: 455). Therefore, war terminology can be used to talk about an argument as in the expression: *"His criticisms were right on the target."* (Lakoff & Johnson 1980: 455- 456). In addition to these types of structural metaphors, Lakoff and Johnson (1980: 462) also present orientational metaphors that arise from our spatial experiences, such as *good is up* (e.g. *He's at the peak of health)*, and *bad is down* (e.g. *He came down with the flu)*. Thus, if the learner is familiar with metaphorical language in general, and the CM(s) in the unfamiliar idiom, comprehension may be supported. Yet not everyone is aware of conceptual metaphors in language (Lakoff & Johnson 1980: 455), which reflect the cultural-specific conceptions (Lakoff & Johnson 1980: 465-467). Also, this strategy can only be used with semi-transparent idioms, not more opaque ones. In addition to metaphors, figurative expressions can also include metonymy. According to Kövecses and Szabó (1996: 338), metonymy involves "a stand for conceptual relationship between two entities within a single domain". To illustrate this, frequently used metonymies involving the human *hand* include *hand stands for person/activity/skill* (Kövecses & Szabó 1996: 341-344).

Research on the use of CMs has focused on the implementation and testing of various pedagogical techniques, such as grouping idioms under CMs. The results of these studies point to the superiority of the CM strategy (Kövecses & Szabó 1996; Beréndi et.al. 2008). The results of a few studies on the autonomous use of strategies, nevertheless, paint a different picture. For instance, Beréndi et.al. (2008), in one of their experiments, discovered that when instruction on CM was minimal and more implicit, no benefits for learning were found (Beréndi et.al.

2008). Furthermore, Skoufaki (2005) found that Greek EFL-learners rarely used knowledge of CMs when guessing the meanings of unfamiliar idioms in or without context. In contrast, Katsarou (2010) observed much higher rates of metaphor use for guessing the meaning of idioms. Nevertheless, the use of metaphor did not correlate with comprehension. In summary, these findings suggest that the successful use of CM in an autonomous setting is unlikely.

Grant and Bauer (2004: 51) discuss the use of metaphors from a slightly different perspective, arguing that figuratives, as a distinct category from idioms, have a recognizable untruth that can be "reinterpreted pragmatically to understand the intended truth". In other words, according to Grant and Bauer, figurative language can be undone to work out the meaning of the figurative. This is illustrated with the expression "*He's gone off the rails"* = "he's no longer on the rails" = "he's not moving forward in a controlled guided way" (Grant and Bauer 2004: 51). In *the Collins Cobuild dictionary* the idiom *go off the rails* implies not just "wild" and dangerous behavior, but also doing something that is unacceptable and upsets other people. When Grant and Bauer's "unpicking" of the idiom and the dictionary definition are compared, it is clear that the literal and the figurative meaning share many features, which may help in comprehension. Yet small nuances of the meaning can easily be disregarded, and misinterpretations may occur if, for instance, the literal meaning is overgeneralized.

Building on the work of Grant and Bauer (2004), Nation (2013: 506) proposes an interpretive strategy for learning of figuratives (see also Grant & Nation 2006: 9). It involves (1) understanding the figurative meaning, (2) understanding the literal meaning, and (3) bringing the two meanings together. This strategy builds on the presumption that the learner already has access to both the figurative and the literal meaning. Hence it is a strategy for understanding the motivation behind the literal meaning. A similar strategy where literal meaning was used as a key to metaphorical meaning was found by Cooper (1999: 249). In the case of some opaque idioms, the figuration can also be explained by assigning roles to words or phrases in the idiom, such as beans standing for secrets in *spill the beans* (Boers & Lindstromberg 2009: 14). In this study, all of the above-mentioned strategies, in which metaphor is used in one way or another, will be considered as clues used for guessing strategies.

All three strategies overlap to certain degrees. Pictures, in particular, can depict both the etymology or conceptual metaphors and metonymies. Still, all of the strategies have their specific characteristics by which they can be distinguished from one another. The literature reviewed here for the most part introduces complex strategies that involve multiple steps and require awareness of idiom characteristics, training, assistance from a teacher, and skills in using the strategy. Therefore, it can be questioned to which extent, Finnish upper secondary

students will have access to these strategies when discovering the meaning of unfamiliar idioms. Moreover, the usefulness of these strategies for learners may depend on the learners' cognitive style (e.g. Boers et.al. 2008). However, when combined with other sources of information, these strategies may help the learners see the motivation behind the literal meaning of the idiom, which in turn will make learning more effective.

In this study, I aim at answering the following research questions: (1) What vocabulary learning strategies do Finnish upper secondary students use in discovering the meaning of English and Swedish idioms while thinking aloud? (2) How well are students able to discover the meaning of English and Swedish idioms? What types of errors do they make? and (3) Are there any differences between discovering the meaning English and Swedish idioms? These research questions were chosen since very little is currently known about the VLS that Finnish language learners use. Moreover, whereas idiom comprehension has been researched in the Finnish context (Mäntylä 2004), no previous research has been conducted on the extent to which learners are able to discover the meaning of idioms when they are allowed to use a broader range of aids, such as a laptop. Lastly, no studies were found on VLS and idioms in which the task performances regarding two different languages would have been compared.

7. DATA AND METHODS

Data was gathered with a concurrent think-aloud and a screen capture video that recorded the participant's engagement in a think-aloud task. In this section, I will begin with describing the procedure for selecting English and Swedish idioms for this study (7.1). I will then move on to describe the methods for data collection in more detail (7.2). After this I am going to discuss ethical considerations (7.3), and lastly, present the method of analysis (7.4).

7.1 Selecting idioms for this study

I chose two idiom dictionaries from which I selected the idioms used in this study. The English idioms were chosen from the fourth edition of the *Collins Cobuild* (2020) *Idioms Dictionary* and the Swedish idioms were chosen from Hans Luthman's (2017) *Svenska idiom: 5.000 vardagsuttryck.* The main reasons for choosing these two dictionaries was that they were both rather recently published when compared to other Swedish and English dictionaries on the market. Relevancy is crucial since language is constantly evolving and new multi-word items come into play while old ones drop out (Schmitt 2010: 158). Moreover, the theory of idioms has developed throughout the years, and research on idioms has introduced new insights, both of which are of great value to dictionary publishers. Another criterion for choosing these two dictionaries is that they both provide example sentences, which were used in this study to elicit the guessing from context strategy. Some difficult words in the dictionary sentences were replaced with more frequent synonyms, so that guessing would be easier.

In its introduction, the *Collins Cobuild Idioms Dictionary* provides a somewhat comprehensive account on idioms and the different information the dictionary supplies the reader with. In the definition of an idiom, semantic opacity is covered. This is revealed by the definition that can be summarized as follows: idiom is a metaphorical phrase whose meaning is different from the meanings of the single words it consists of (Collins Cobuild 2020: v). Additionally, the variability regarding the easiness of deriving the metaphorical meaning from the literal meaning is discussed. Lexicogrammatical invariability or "fixedness" is questioned as a feature of idioms, and this is supported by the evidence found from the Collins Corpus. A wide array of idiom subcategories are included in the dictionary, as listed in the introduction: traditional idioms, multi-word metaphors, similes, proverbs, pragmatic expressions and

expressions that are combinations of phrasal verbs (Collins Cobuild 2020: v-vi). This highlights the broadness of what mwus are counted as idioms in the dictionary. As for the dictionary entries, multiple aspects of form, meaning, and use are provided. These include the definition of the idiom, style, and context of use (incl. dialect/region of use), two authentic contextual sentences taken from the Collins Corpus, idiom variations, synonyms, and etymology.

The Swedish dictionary is substantially less comprehensive. First, it has to be noted that the dictionary is not based on robust corpus data like the *Collins Cobuild dictionary* but rather on the author's observations and data collected from other unnamed dictionaries. However, native speaker-judgment has also been recognized as one way of identifying formulaic sequences (Wood 2020: 39-40). Probably the most significant deficiency in the dictionary is its lack of an idiom definition. The only definition that can be deduced is juxtaposing "everyday phrases" with idioms. This would suggest that idioms consist of multiple words and are often part of informal language use. As for transparency of meaning, the mwus included in the dictionary vary a lot. In the introduction manual, phrasal verbs are singled out as one subcategory of idioms. They are marked with an accent on the particle because in Swedish the particle of the phrasal verb is always stressed. Additionally, similarly to the *Collins Cobuild dictionary*, proverbs, similes, and lexical phrases can be found in the dictionary. The dictionary entries are quite narrow and include only one or two definitions of the idiom, synonymous idioms, and an example sentence. Furthermore, it is not explicitly stated whether the example sentence is authentic, slightly modified, corpus-inspired or constructed for dictionary purposes (see Hult 2009: 220).

One of the criteria for choosing idioms for this study was frequency since it is the most important criterion when choosing vocabulary items to learn. The higher frequency the words have, the more likely they will be encountered in language, which makes them more valuable for learning (Webb & Nation 2017). In this study, participants had to discover the meaning of unfamiliar idioms, and hence, frequency does not have as pivotal role as it would if existing idiom knowledge was investigated. However, in order to make the task as authentic as possible, frequency was taken into account. Frequency may also affect the extent and accessibility of information on the internet. The *Collins Cobuild Dictionary* had marked roughly 1,200 most frequent idioms according to their frequency in the Collins Corpus. The first stage of selection was to compile a list of them. The Swedish idiom dictionary did not contain frequency markings, so I chose to use Google as a corpus to determine frequency (see also Prentice 2010; Wood 2015: 22). The borderline for selecting an idiom was 20 000 hits at the time of googling.

Interestingly, some of the idioms received only a few hundred hits, which highlights the importance of taking frequency into account.

After determining the frequency, random idioms were chosen in a systematic order (e.g. every 20th idiom in the list). Of these idioms, firstly, the subcategories of idioms that were excluded from the study were eliminated (see 4.2). At this point, I also chose to exclude English idioms that were specific for British, American, or Australian English, since learning these is not ideal due to the limited scope of use. Additionally, English and Swedish idioms that contained derogatory or inappropriate language were excluded. Secondly, with the help of my own knowledge of Finnish idioms, those idioms that had a full or partial equivalent Finnish idiom in form and/or meaning were eliminated (see 4.1). Finally, idioms that were least transparent and non-literal were chosen for this study. Altogether 23 idioms for both languages were chosen for this study.

7.2 Data collection

In the following section, the participants in this study will be described. After this the think-aloud method, the think-aloud task procedure, and the video screen capture method will be presented. Lastly, I am going to discuss some ethical issues and provide an account for the method of analysis.

7.2.1 Participants

Altogether nine upper secondary school students were recruited for this study. Upper secondary students were chosen for the study since they already have plenty of experience in language and vocabulary learning. In addition, upper secondary students have higher metalinguistic awareness than for instance primary school students. All of the participants were second-year students. Five of them were female and four were male. They were recruited from one upper secondary school located in Central Finland. The criterion for selection was an intention to complete the matriculation examination test in both English and Swedish, so that students who may not be interested in the two languages and therefore may have a lower-than-average proficiency in them, particularly Swedish, would be ruled out from the study. This can nonetheless not be guaranteed. None of the eight second-year students had taken the matriculation exam in English and Swedish. All participants rated their course average and perceived language proficiency level to be eight or higher on a scale of four to ten. Hence, the

participants seemed to be on similar proficiency levels in both languages, but it is most likely that their English proficiency was higher than their Swedish proficiency, since the majority of Finnish students start studying English earlier than Swedish. In sum, the participants outwardly seem to form a rather homogenous group.

In the recruiting process, the principal was first contacted via email to gain permission for doing research. After receiving approval from the principal, language teachers were contacted to recruit students. The teachers who wanted to collaborate, gave permission to briefly introduce the study and collect the contact information of students that wanted to participate. These students were thereafter contacted individually in order to arrange for a meeting. Data was gathered during March and April 2022 at a peaceful place where there were no interruptions or disturbances. After data collection, the participants received a free cinema ticket, and task feedback as a reward for their time and effort. Additionally, the participants who wished so, were sent a file on vocabulary learning strategies via email.

7.2.2 Think-aloud

Thinking aloud (TA) or verbal report is a voluntary activity in which learners talk their thoughts out loud while they are engaging in a relevant task, and a think-aloud protocol is a recording of that reporting (Cowan 2017). Thinking aloud is useful for uncovering "the cognitive processes of human information processing" (Zhang & Zhang 2019: 304). The method is used particularly in the field of psychology and cognitive science, but also in several subfields of applied linguistics (Zhang & Zhang 2019: 302). As vocabulary strategies are at the core of this study, TA is a suitable method since it allows the analysis of some otherwise unobservable behaviors or processes related to language learning (Zhang & Zhang 2019: 305).

The TA method has been used in studies investigating vocabulary strategies (e.g. Gu 2003a; Gu et.al. 2005; Lawson & Hogben 1996), reading strategies, particularly inferring meaning from context (e.g. Bengeleil & Paribakht 2004; Hu & Nassaji 2014; Nassaji 2003, 2006), and idiom comprehension strategies (e.g. Cooper 1999; Wray et.al. 2016). Nation (2013: 519-520) also states that thinking aloud can be used in testing the declarative knowledge, i.e. what the learner can do, of several vocabulary strategies such as guessing from context. Moreover, the TA method has been used to gather data of dictionary use, hence that participants verbalize their thoughts while engaging in dictionary activities. These studies have investigated dictionary use in relation to various domains of L2 learning, such as L2 reading (e.g. Peters 2007), and L2 error correction (e.g. Wolfer et.al. 2018). In the studies of Peters (2007) and Wolfer et.al. (2018), screen capture technologies were also made use of alongside TA. Similar

to this study, one group in Chan's study (2012), had to think aloud while interpreting lexical items with the help of a dictionary. Based on the methods employed in these previous studies, it can be concluded that TA is a well-established method that fits the purposes of this study well.

There are two different kinds of TAs. Concurrent TA refers to thinking aloud while performing a task (Ericsson & Simon 1993: xiii), whereas in retrospective TA, the participant is asked to think aloud about an activity that has already been performed (Zhang & Zhang 2019: 303). In this study, concurrent TA was used for several compelling reasons. Firstly, in retrospective think-aloud, thoughts that are stored in the short-term memory may have been forgotten unless the task is very short (Ericsson & Simon 1993: xvi). In other words, by using concurrent think-aloud, the risk of forgetting is minimalized. Secondly, during the time between task completion and retrospective TA, the participants may have elaborated the original cognitive process and information, for instance by including rationalizations and justifications (Ericsson & Simon 1993: xvi). Consequently, their verbalizations may not be accurate descriptions of their thinking that occurred while completing the task. Thirdly, concurrent TA is more widely preferred since it is methodologically more valid (Fox et.al. 2011: 317; Zhang & Zhang 2019).

Cohen (2013) distinguishes between *self-report*, *self-observation,* and *self-revelation*. Firstly, in self-report, learners describe what they usually do. Secondly, self-observation is the inspection of specific language behavior, and it can be done both intro- and retrospectively. Thirdly, self-revelation fits in the above definition of concurrent TA. In this study, self-revelation will be the main source of data. However, the action-oriented TA task may also prompt self-observation, since the participants cannot be expected to produce ongoing TA while solving complex linguistic problems. As a result, participants may make retrospective observations about what they did or thought, but within a short time frame to the language behavior that they manifested so that the information will still be present in short-term memory (Ericsson & Simon 1993: xvi). This can be ensured with the use of prompts (see 7.2.3).

Ericsson and Simon (1993) have also proposed three different levels or types of verbalizations. Type 1 verbalizations are direct reproductions of information in the form in which it was heeded (Ericsson & Simon 1993: 16). Type 2 verbalizations involve explication of the thought content (Ericsson & Simon 1993: 79). In other words, the subject verbally codes already existing information. Type 3 verbalizations are interpretative in nature, and they include explanations of thought processes or thoughts (Ericsson & Simon 1993: 79). Ericsson and Simon (1993: xiv) distinguish social verbalizations (Type 3) such as explanations,

descriptions, justifications, and rationalizations from thinking aloud since these require additional processes and may also change the natural sequence of thoughts. Contrastively, in thinking aloud (Type 1 and Type 2) "subjects verbalize new thoughts and generate intermediate products as these enter attention" (Ericsson & Simon 1993: xiii). The main argument of Ericsson and Simon (1993) is that only Type 3 verbalizations influence and change the cognitive processes and affect task performance. For this reason, the function of verbalization should be made clear to the subject when giving instructions (Ericsson & Simon 1993: xiv) and using probes during the task. In this study, the aim was to elicit Type 1 and Type 2 verbalizations.

The rate of verbalizations may vary and sometimes the TA may be incomplete, for example, if there are longer pauses in the TA protocol or the information is uninformative. Based on several early studies, Ericsson and Simon (1993: 252) provide a categorization of the reasons that may lead to incomplete concurrent reporting. Their analysis shows that incomplete reporting could occur (1) while reading and trying to comprehend the text, (2) during intense cognitive activity such as problem-solving, and (3) during mediating steps that lead to a solution. Based on these observations, it is presumable that the participants in this study will not be able to continuously maintain a high level of thinking aloud, since the task may steer the participants into the above-mentioned situations.

When inspecting the validity of the TA method, aspects of reactivity and veridicality should be considered. The question of reactivity relates to whether the additional load of having to think aloud changes cognition (Zhang & Zhang 2019: 305). The underlying question of veridicality in turn relates to whether the data accurately reflects the cognitive processes of the participant (Zhang & Zhang 2019: 306). Although both reactivity and veridicality of TAs have been questioned by some scholars (e.g. Barkaoui 2011), Hu and Gao's (2017: 186) review of self-regulated reading studies adopting TA indicate that two different strategies, when carefully implemented, can be used to minimize reactivity effects maximizing veridicality. These two strategies are training and prompting, which will also be used in this study (see 7.2.3). Both of these practices are widely used and recommended practices within the TA research since they increase the validity of the study (Ericsson & Simon 1993: 253-257**;** Hu & Gao 2017: 186; Zhang & Zhang 2019: 307). Moreover, as stated earlier, the choice of using concurrent TA will diminish the concern of veridicality (Zhang & Zhang 2019). In addition to considering these aspects, the choices related to the type of TA and language can be controlled in order to increase the validity of the data.

A meta-analysis done by Fox, Ericsson, and Best (2011) of 94 independent data sets including nearly 3,500 participants altogether, revealed no reactivity effects for thinking aloud when compared with silent conditions. However, if the participants' thoughts were directed, for instance by asking to provide explanations (Type 3), the think-aloud altered the performance by improving it. The study strongly supports Ericsson and Simon's (1993) claims, according to which non-metacognitive verbalizations of thoughts (Type 1 and 2) do not affect the accuracy of task performance. However, the meta-analysis showed that verbal reporting does tend to prolong task performance. In sum, although some critical issues have been raised about reactivity effects, Fox et.al. (2011) study proves that it is not a major concern as long as the type of TA is carefully controlled.

It is also by no means indifferent, what language or languages the participants are allowed to use when thinking aloud. Zhang and Zhang (2019: 306) report that many researchers have allowed the participants to speak in the language of their choice, accepting also codeswitching. In this study, the participants were allowed to speak in their L1, Finnish, so that language proficiency would not hamper or restrict thinking aloud (Cohen 2013; Smith & King 2013: 717). The use of the target language (English or Swedish) was also permitted. All in all, with different choices that decrease reactivity and increase veridicality, TA can be considered as a well-established and reliable method for collecting rich data about the participant's cognitive processes (Zhang & Zhang 2019).

7.2.3 The think-aloud task procedure

In this study, a think-aloud task, similar to the one Lawson & Hogben (1996) used in their study of Australian students' VLS, was used to collect data. The task was to discover the meanings of unfamiliar English and Swedish idioms. A vocabulary learning task such as this would give information about what learners are able to do, instead of solely relying on learners' reports (Nation 2013: 334). Before the actual think-aloud task, a pilot study with three participants was conducted. This gave me an opportunity to make changes in the task setting, the instructions, and the methodology. The changes related to my role as a researcher and more specifically, the types of prompts I used to encourage verbalizations. In addition, idioms with translation equivalents were excluded, and video recordings were abandoned as a means of collecting data. Moreover, the time was limited to 20 minutes per language (40 minutes altogether) so that the task would not be too long and exhausting.

The data collection consisted of the following steps: 1) informing about ethical practices, 2) signing consent forms and filling in a short background information survey (see

Appendices 3 and 4), 3) TA training, 4) choosing unfamiliar idioms, 5) task instructions and, 6) completing the task (for more detailed task instructions see Appendices 1 and 2). The TA training in the form of a quick practice task, was implemented to ensure that the participants would be introduced to verbal reporting (Zhang & Zhang 2019). In the practice task, the participants were asked to verbalize their thoughts on an ambiguous image with the pseudoword *steck* written below it. The practice task differed from the real task so that it would not have an effect on it. In addition to the other benefits of training, it turned out to give valuable information about the participants' ability to verbalize, which guided the quality and quantity of prompts during the actual task performance.

Before accomplishing the real task, a simple yes/no test or self-assessment was completed to ensure that the target idioms were unknown to the participant in question. In the yes/no test, the form of the idiom was shown, and the participants had to indicate whether they knew the meaning of the idiom or not. At this stage the example sentences were not yet shown. If a target idiom was recognized as a fully or partially known idiom, it was replaced by a backup idiom. Idioms are low-frequency items and for this reason it turned out to be highly unlikely that the participant knew the meaning of idioms. Since the participants were quick to judge the idiom as known if they were familiar with the single words the idiom consisted of, participants were directed to think if they have seen the words together and know their meaning as a whole. Furthermore, some participants very confidently stated yes to many idioms, and in order to not run out of idioms, these students were asked to explain the meaning, a practice that can has been used in pretests (e.g. Peters 2016: 120). If the meaning was somewhat correct the idiom was disregarded, and if it was wrong, the idiom was chosen. Instead of saying yes/no some students also stated that they have seen the idiom but have no idea of the meaning. Albeit this may be helpful in discovering the meaning of the idiom, the idiom was classified as unknown. Yes/no tests have been used for vocabulary size, and vocabulary placement tests (Nation 2013: 539-540), nonetheless, in this study, no general conclusions will be drawn of the participants' vocabulary level. The fact that false positive answers will have no effect on the task supported the choice of this test. After the test, the participants received cards that contained the chosen unfamiliar idiom and an example sentence.

After selecting seven English and seven Swedish unfamiliar idioms, the participants were given task instructions in Finnish. In brief, the participants were instructed to discover the meaning of the chosen phrases with the help of any available means while thinking aloud. During the task, the participants had access to idiom cards, and a laptop. They were also informed about the possibility to use their phones, and any other means they would find helpful

in completing the task. Additionally, the participants were instructed to fill in a simple answer sheet in which they had to write the idiom and its meaning in Finnish, English or Swedish. The participants were also instructed not to write the sentence on the sheet. Five of the participants started the task with Swedish idioms, and four with English idioms, in order to reduce any language effects.

During the TA task, if the participant stopped verbalizing, they were encouraged to keep thinking aloud in order to ensure the richness of data. Various prompts, such as "keep talking", "What are you thinking now?" (Ericsson xiv), and "keep reading aloud" were used. The prompts were as neutral as possible, and they did not take the form of specific questions. Prompts were not used when the participants typed on the keyboard or wrote on the answer sheet. Once the task was completed, participants were thanked for their participation and given the opportunity to ask questions.

The participants did not know that the study focuses on idioms, but rather on "vocabulary" and "phrases". Vocabulary learning strategies were either not mentioned, and the participants were only told that the aim of the research was to merely observe vocabulary learning. If the participants were recruited from the same course, they were instructed not to say anything about the data collection procedure to their peers who had also signed up to participate but had not yet met with the researcher. The reason for not revealing the specific purpose of the study, and not describing the data collection procedure in advance was to minimize the reactivity effects (Alasuutari 2007: 97), and the effects this knowledge may have had on task performance. However, as reported earlier, the participants may have had some indication that the words in the idiom tend to occur together, and have a unified meaning based on the instructions. Furthermore, in the few cases where the participant had erroneously thought that s/he knew the meaning of the idiom, which was still judged as unknown by the researcher, the participant may have figured out that the phrases cannot be interpreted literally. The points raised here demonstrate the difficulties associated with selecting unfamiliar idioms for a vocabulary learning task. The data collection lasted altogether from 45 minutes to an hour, depending how quickly the participant performed the task.

7.2.4 Video screen capture

TA protocols can be integrated with other sources of data, such as observations (Hu & Gao 2017: 188). Zhang and Zhang (2019) advocate for methodological triangulation, as one possible solution for the drawbacks of the TA method. Due to the possible incompleteness of

the TA and the complexity of the use of online sources, such as dictionaries, video screen capture was chosen to complement TA data. Video screen capture has become popular in instruction manuals and tutorials and also more recently, as a useful tool in research and education (Hamel & Séror 2016: 140).

Thanks to new technological tools, observing computer-mediated language learning practices have recently become more convenient (Hamel & Séror 2016). Video screen capture (VSC) is a type of software that "allows one to record a movie of on-screen actions, which occur as an individual interacts with a computer" (Hamel & Séror 2016: 138). Earlier, computer-based tracking devices could be divided into computer logs and recording software, the former being text files in which mouse clicks, keypresses, and entries in text fields are captured, and the latter being videos of any activity visible in the computer monitor (Fischer 2007: 411). This distinction remains relevant even today, however, the two types of software can also be used in conjunction with one another (Hamel & Séror 2016: 141). Alternatively, some software may allow customization of the sources of video input (Hamel & Séror 2016: 139). In this study, all on-screen activities and sounds (TA) were recorded with VSC.

Tracking tools, such as log files, have been utilized in several studies on dictionary use. The research in this branch can be viewed from two separate but overlapping perspectives: the dictionary user and the language learner. Studies conducted from the perspective of the dictionary user have focused on how "usable" and efficient the different features of the dictionary are (De Schryver 2006; Hamel 2012; Verlinde & Binon 2010). The information from these studies is particularly valuable for the dictionary compilers who can use it to further improve the design of the dictionary interface. Studies conducted from the language learner perspective have looked at the value of dictionary look-up for diverse aspects of L2 learning (e.g. Chen 2016; Laufer & Hill 2000; Lew & Doroszewska 2009; Peters 2007). Tracking technology has also been used to record other kinds of online activities. For instance, screen capture technologies have been used in two L2 writing studies in which participants had access to various online sources (Gánem-Gutiérrez & Gilmore 2021; Yoon 2016).

VSC has many advantages compared to other methods that have been used in dictionary studies. Firstly, Szczepaniak (2006: 86) notes that screen capture video is a more reliable and effortless way of keeping track of what the students actually do with a dictionary, meanwhile filling in self-reports alongside dictionary use can be time-consuming and burdensome for the participant. Secondly, interviews and questionnaires may only reveal the users' subjective perceptions of dictionary consultation (Welker 2006: 26 as quoted in Tarp 2009: 284). Additionally, the limitations of retrospective reports (see 7.2.2) are also of relevance with

surveys and interviews, unless the goal is to investigate dictionary use habits. Thirdly, relying solely on the researcher's observations or field notes involves the risk of disturbances (Tarp 2009: 286) and observer bias, which are not a concern with VSC since it can capture actions in more detail and high quality. Thereafter, the video can be reviewed if necessary. Lastly, Hamel (2012: 358) notes that using screen-capturing to record learner-task-dictionary interaction "proved to be a direct, objective and non-intrusive method of collecting dynamically empirical data". In conclusion, the arguments presented above advocate for the use of VSC in this study.

Despite the several advantages of VSC, it also has limitations. Log files that register the transactions between the user's computer and the dictionary database, can be collected automatically without the presence of a researcher. However, due to the learning focus and the use of TA, this study calls for the presence of a researcher. Thus, it is possible that the presence of a researcher and the awareness of the VSC may have changed the behavior of the participants and created an observer's paradox. Moreover, Tarp (2009: 290) notes that log files "only furnish data related to the external aspects of the consultation process" dismissing the inner aspects, such as motivation, needs, and results of the consultation. For this reason, log files should be accompanied by other methods (Tarp 2009: 290). In this study, TA was used to gather data about other strategies than those that require a computer and to create a deeper understanding of the different strategies used on the computer.

During the idiom comprehension task, VSC software called SCRE.IO was used to record speech and capture any on-screen activities. The recorder did not cause any disturbances. Additionally, all previous browsing history from the laptop was deleted so that it would not direct participants, for example, to use certain dictionaries or keywords. Participants did not have to be trained to use the VSC, they were only informed about it. Unfortunately, two videos were lost for technical reasons, and therefore could not be used. For these participants, only the idiom answer sheet was analyzed. Therefore, the data consists of altogether seven VSC recordings that contain the TA protocol and nine answer sheets.

7.3 Ethical considerations

There were many ethical aspects that had to be considered when conducting this study. In general, throughout the whole research process, the TENK guidelines of the Finnish National Board on Research Integrity (2019) were applied. Participants were recruited via upper secondary schools. Before gathering any data, the permission of the principle, and language

teachers were obtained. Before data collection, the participants were briefly verbally informed about the methods and handling of data, the rights of the participant, and crucial ethical considerations, such as anonymity. After this, written consent for participation was collected. Moreover, after data collection, a data privacy announcement was sent to the participants via email (see Appendix 5). It included information about the intent of the study, the rights of the participant, and the storage, handling, and deletion of personal data. When reporting the results of the study, pseudonyms were used and any identifiable information was excluded, to guarantee anonymity of the participants.

Initially, participants were attempted to recruit without a reward, being aware of its drawbacks. However, although some participants volunteered, they canceled their participation, perhaps due to the length of the data collection. As a consequence, a movie ticket was used as a reward for participation. It was also mentioned that participants would receive feedback on the vocabulary task, and that participation could be beneficial for vocabulary learning. Together these turned out to be a sufficient inducement for participation. Nevertheless, in the desire for a reward, some unmotivated participants may have volunteered, which may have led to negligent completion of the task. Furthermore, it is likely that the study may have attracted students, who are motivated language learners, and particularly interested in enhancing their vocabulary skills.

7.4 Method of analysis

In this study, content analysis was adopted as a method of analysis. It is a basic method for interpreting qualitative data that is in textual form (Tuomi & Sarajärvi 2018: 103, 117). Qualitative content analysis is "a powerful analytical method used for the subjective interpretation of the contents of both qualitative and quantitative steps in a systematic and context-dependent manner" (Marying 2014; Schreier 2014 as cited in Selvi 2019: 442). Since the method can be used for both qualitative and quantitative analyses, it is sometimes considered as a mixed methods approach (Mayring 2014: 10; Tuomi & Sarajärvi 2018: 138). The features characteristic for the method include systematicity, flexibility and reducing the data (Schreier 2014: 2). Content analysis was chosen due to its suitability with the data collection methods used, and the diverse opportunities it offers to compare, describe, simplify, and summarize extensive and multifaceted data of a somewhat novel topic of research.

The content analysis used in this study had both inductive and deductive features (see also Tuomi & Sarajärvi 2018:110-111: *teoriaohjaava analyysi*). It was therefore guided by earlier theoretical frameworks and research findings, while allowing the emergence of new categories purely based on observations made from the data (Tuomi & Sarajärvi 2018: 109). In Schreier's (2014: 9) terms, categories were both data- and concept-driven. In sum, the method of analysis used in this study combines and alternates between inductive and deductive approaches, however leaning more towards deductive analysis since theory was incorporated in the analysis right from the start (Tuomi & Sarajärvi 2018: 110, 113). For example, strategy categories were mostly selected from existing strategy taxonomies and strategy research.

Before any analysis could be done, the TA protocols and VSC were transcribed verbatim into written form. The transcriptions consisted of speech and descriptions of what was said, and what on-screen activities were executed on the laptop. Altogether the transcriptions yielded roughly 16 000 words of text. As an example of the transcription, dictionary use was described as following: search 1 *can of worms* in sanakirja.org. In order to register what crucial information was left unnoticed, some information that only appeared on the screen was also described in the transcription. The answers of the idiom task were also written on another separate document. These transcriptions enabled later coding and classifications. Next, criteria for scoring the answers were established. In addition to fully correct answers, partially correct answers knowledge were also credited. The criteria for scoring the answers were the following:

2p - the answer matches or resembles the dictionary definitions provided by *Collins Cobuild dictionary* and/or *Oxford idioms dictionary*, or *Svenska idiom*. The central meaning of the idiom has been understood

1p - some part of the answer is correct but other part is incorrect (e.g. literal or overgeneralized from context), the answer does not completely match with the correct definition; some essential meaning is missing or the meaning is too specific or too vague.

0p - the answer does not resemble the meaning of the idiom at all (e.g. literal translations)

Ideally, to enhance the reliability of the scoring, two raters would have independently scored the answers according to same criteria, and then interrater agreement would have been calculated. However, due to the small scale of this study, the answers were initially scored only by me. Those answers, which scoring was indecisive, were resolved through discussions with

my two supervisors. For the answers that were completely (0p) or partially (1p) erroneous, a coding scheme was created based on the literality of the answers and the source of the error. It consisted of four main categories of which the last also has three subcategories. The categories are presented in table 3 below:

Table 3. Error categories

Error type	Description
1) Literal	word-by-word translations of the idiom or some of its constituent words; wrong senses of words were also counted
2) Juxtaposition	juxtaposing the meaning of the idiom with part of the example sentence
3) Literal + non-literal	a definition in which there is a literal and a non-literal part
4) Non-literal	a) non-literal meaning that fits well in the example sentence and is inferred from context b) Finnish idioms; often have formal or semantic similarities with the target idiom c) other non-literal answers; often dictionary definitions that were erroneously interpreted or incomplete

The categorization is not perfect, since at times the sources of errors may have overlapped. Additionally, category 4c contains miscellaneous answers that did not fit in any other category. However, the common denominator of most of the answers in category 4c, was evidence of having consulted a dictionary definition.

A coding scheme for the strategies was also created based on data and previous research. It is described in table 4 below:

Table 4. Strategy taxonomy and criteria

Strategy	**Criteria**
Idiom selection	Ordering the idioms or suspending the meaning discovery of an idiom.
Assessing knowledge of a word or an mwu	Explicitly commenting about how well and to what extent a word or an mwu is known. Questions, such as “what does this mean”, were not counted, since they were more implicit. Evaluations of

	comprehension were not counted, since they would have been extremely difficult to distinguish from similar informal comments such as: “umm”, “I don’t know”, “well that didn’t help much” and so forth.
Guessing strategies	Making a guess about the meaning of a word or an mwu. Consists of two subcategories: (1) guesses before dictionary consultation, and (2) guesses after dictionary consultation. Both categories can involve guessing from context. Literal translations, reading aloud or translating dictionary definitions, and repeating an earlier guess were not counted.
Using a search engine	Searching words or mwus with Google or other search engine. Accessing a particular webpage with Google was not counted. If a word included in the search was instantly corrected, or a suggestion (did you mean) was clicked on, it was counted as one search. If a new word was added to a previous search, it was also counted as a new search.
Monolingual dictionary	Using a monolingual dictionary. Dictionary look-ups, i.e. searching a word or mwu in a dictionary, and reading aloud dictionary information were counted. If new words were added, words were changed, or words were removed from the previous dictionary search, these were counted as separate look-ups. Also, returning to a dictionary website and re-reading dictionary information was counted as a separate instance of the strategy, whereas reading multiple parts of the same dictionary entry was counted as one look-up. If an inflected form of a word or phrase was looked up, and instantly, a suggestion was clicked on, this was regarded as a single look-up. The reason for this is that, if all the erroneous or inflected searches had been counted, the results would have been distorted.
Bilingual dictionary	Using a dictionary that had information both in L1 and L2. Changing the target language of the dictionary was counted as a separate look-up. Other criteria are the same as for monolingual dictionary.
Using other webpages	Opening or reading any webpage that is not a dictionary-like webpage.
Using English	Using English for discovering the meaning of Swedish idioms. Instances of Google search containing references to English, and dictionary use that involved English were counted as examples of this strategy. Hence it overlaps with both using a search engine and

	using a dictionary. Other comments about the use of English were not counted.
Other strategies	This category included all other leftover strategies that occurred only with one participant or only a minimal number of times not enough to deserve an own category. These strategies include compound word analysis, word part analysis, etymological elaboration, and retention of words in chunks.

Instead of solely relying on qualitative analysis, quantitative analysis was also used, since together they enable arriving at complementary conclusions that together capture the complexities of the phenomena (see Tuomi & Sarajärvi 2018: 120-121). Content analysis can be quantitative when content is quantified or itemized, and then presented numerically or statistically (Tuomi & Sarajärvi 2018: 118-121). The following information was analyzed quantitatively: instances of strategy use, idiom task scores, the type of errors, and approximate task duration per language. Additionally, the number of idioms whose meaning was discovered were calculated. If the participant had finished the discovering and provided an answer for the idiom, it was counted as one idiom. When the participant did not have time to finish the discovery of an idiom regardless of whether they provided an answer for the idiom or not, it was counted as half an idiom. The results of these quantitative analyses were presented in numbers, figures, and mean scores. This allowed to make comparison between English and Swedish.

8. STRATEGIES FOR DISCOVERING THE MEANING OF IDIOMS

The participants had different approaches to the task, and they used different types of strategies in varying quantities. In what follows, each strategy type will be described in the order in which it usually occurred in the meaning discovery process.

8.1 Selecting idioms

In this study, the idioms which the participants had to learn were chosen by the researcher, but the participants were free to choose in which order they would discover the meanings of those idioms. This strategy manifested itself in two ways: putting the idioms into a particular order or suspending the meaning discovery of a particular idiom perhaps with an intent to return to it later. These strategies go into Nation's (2013: 328-329) strategy category of *planning*, and more specifically, *choosing words*. In other taxonomies, these types of strategies have been labeled under "selective attention" (Gu & Johnson 1996), "metacognitive strategies" (Schmitt 1997), and "management strategies" (Fan 2003). By using this strategy, some of the learners showed that they were able to determine a goal, for example being fast or effective, and then focus attention accordingly to achieve that goal (see Nation 2013: 329). These learners also demonstrated Gu's (2012: 336-337, as cited in Gu 2019: 282) characteristics of a strategic learner, since they monitored and evaluated what words were worth putting time and effort to.

One participant (B) ordered the idioms according to their perceived difficulty. When asked after the task performance, why this was done, the participant said that he wanted to discover the meaning of as many idioms as possible to maximize the use of time. He had judged the difficulty of idioms based on how many unfamiliar words the phrases and the example sentences included. When he started accomplishing the task, he proceeded from the easiest idiom to the most difficult one since he presumed that the more difficult idioms would take more time. In the instructions, the participants were told that being fast was not the goal but to discover the meanings of idioms to a level of being satisfied and feeling that the idiom is understood. However, with the use of this strategy the participant showed that he was able to prioritize which idioms to maximize efficiency. However, due to a lack of recording, instances of strategy use could not be counted nor described in further detail.

The flip side of this strategy, suspending, was used by two participants (F and G). It entailed leaving a particular idiom unresolved and proceeding to the next idiom. Only six instances of this strategy were recorded; one use by Participant F, and five uses by participant G. The low use of this strategy compared to the use of other strategies is in line with Fan's (2003), and Labontee's findings (2019). Labontee (2019: 119) speculates that a think-aloud task may not have been suitable for eliciting meta-strategic behavior, since the task was a "short-term, low-risk learning experience". Additionally, the task instructions (see Appendices 1 and 2) may not have encouraged the use of this strategy, since both of the participants asked for permission to move on to the next idiom. The findings in this study also suggest that the use of this strategy was specific to certain learners, as was the threshold to using this strategy. For instance, participant F had completed over ten Google searches and dictionary look-ups before resorting to this strategy, whereas two to eight look ups were enough for participant G to proceed to the next idiom. The use of this strategy may also depend on the language since Swedish idioms were suspended more often (5) compared to English ones (1). The somewhat high number of searches preceding the use of this strategy and the struggles experienced by the learners may explain the differences for English and Swedish. The TA reveals that participants experienced difficulties: participant G even explicitly comments on the difficulty of the task twice by saying: "How can this be so difficult", and "Why are these all so difficult?" after multiple dictionary searches. However, suspending was not always caused by difficulties. Participant G motivates the decision to suspend the third Swedish idiom with "putting it on the back burner", perhaps in a wish for some clarity when getting back to it.

The effects of using this strategy may be both negative and positive. Participant G was able to discover the correct meaning of the suspended English idiom, whereas three of the four Swedish idioms were left unanswered, and one was answered incorrectly. In conclusion, the strategy did not seem to help in completing the task. Suspending may also result in forgetting some of the information already gained from dictionaries and guesses. On a more positive note, the use of this strategy may encourage learners to keep on working with the task, since succeeding is important for motivation. Consequently, the small successes with other idioms may in turn help overcome frustration encountered with the more difficult idioms.

8.2 Assessing knowledge of a word or an mwu

Some participants, particularly those who produced a rich TA protocol, explicitly assessed their knowledge of a word or an mwu. These evaluative comments can be divided into no knowledge, and partial knowledge. This finding provides evidence for that words are known to different degrees (Nation 2013: 44), something which some of the learners seemed to be aware of. The comments of having no knowledge of a lexical item contained references to not knowing, difficulty of a lexical item, not understanding, and having no familiarity with the word or an mwu. For example, participant E stated that she had never heard the word *gäller* that occurred in the example sentence, and participant I commented that *have it in for someone* did not sound familiar at all. The comments about partial knowledge mostly concerned familiarity with a word or a phrase, and a recollection of hearing it. It could therefore be reckoned that the participants recognized the form without being able to retain the meaning. For instance, participant H said that *rädd för* 'be afraid of' was a familiar pair of words" but that she had no recollection of what it meant despite remembering hearing it. Figure 1 below shows the use of this strategy:

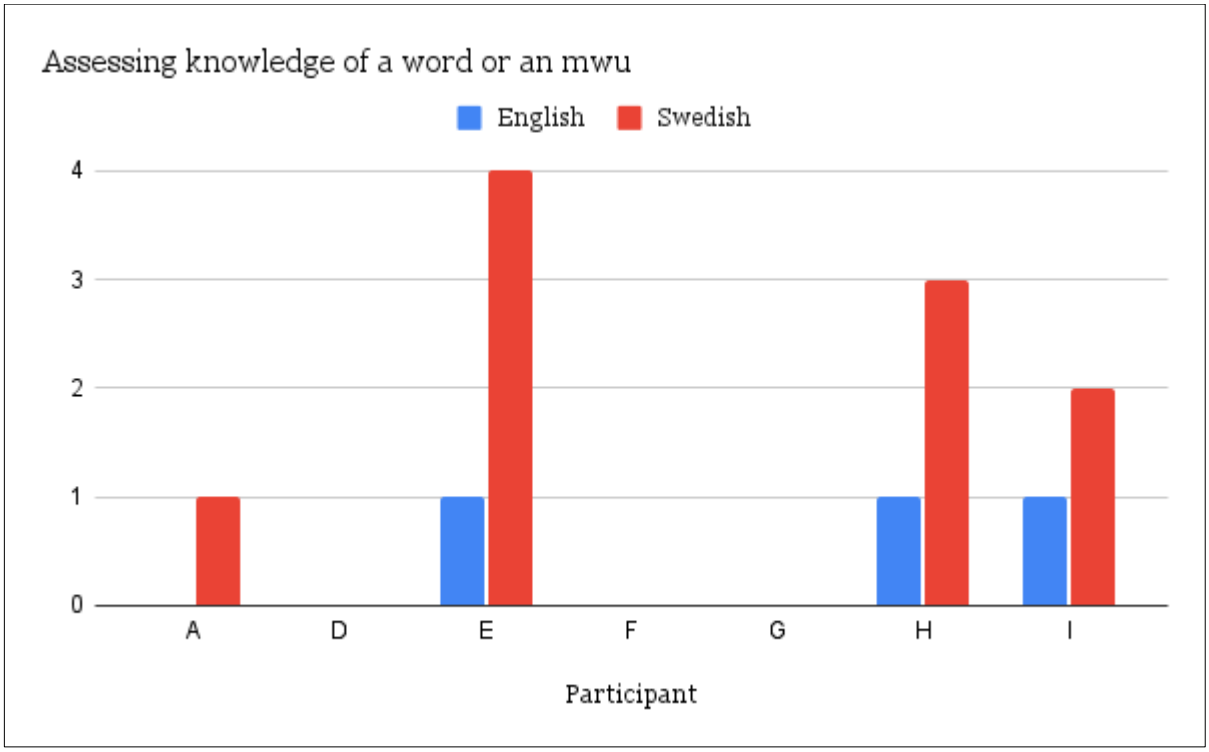

Figure 1. Assessing knowledge of a word or an mwu

This strategy was quite rarely used by four participants, particularly E, H and I. There seems to be a language difference; the meaning discovery of Swedish idioms elicited this strategy altogether ten times compared to the three times regarding English. Since the occurrences of this strategy were quantitatively so low, definitive conclusions should be avoided. The results still seem reasonable considering the hypothetically lower proficiency in Swedish, and hence the larger number of unfamiliar words encountered.

Assessing vocabulary knowledge was counted as a strategy in this study, since Gu (2012: 336-337, as cited in Gu 2019: 282) sees analyzing the task, in this case knowledge of words, as a characteristic of a strategic learner. Moreover, when the meaning of a word or an mwu is not known, assessments of familiarity may help direct and guide further strategic actions. If for instance, a word is fully unfamiliar, multiple comparative dictionaries may be prompted as a result. In contrast, if the form of the word is recognized as familiar, one dictionary search may be enough to help retain or confirm the meaning since the learner may be able to tap into existing implicit vocabulary knowledge. Sometimes, dictionary consultation will not be needed, as in the case of participant H, who soon, with the help of sentence context, recalled that *rädd för* meant to be afraid of something. Other learners may also have engaged in using this strategy but perhaps due to the scarcity of TA, such operations were not visible to the researcher.

8.3 Repetition of idioms

An examination of the data shows that repetition was used as a VLS. This finding is in line with a couple earlier studies in which similar methodology was used (Lawson & Hogben 1996; Barcroft 2009). Other findings of the use of repetition have also been obtained in several self-report studies (Fan 2003; Gu & Johnson 1996; Labontee 2019). However, these results are not fully comparable since repetition in discovering the meaning of a word is not distinguished from repetition for memorization. Yet these studies provide further support for the prominence of repetition as a VLS. The participants in this study repeated multiple lexical items from single words to full sentences. As for the target idioms, multiple kinds of forms were repeated including the idiom embedded in part of the example sentence (*han ville göra rätt för sig*), the canonical form of the idiom (*göra rätt för sig*), idiom variation (making noises), or part of the idiom (*rätt för sig*). Additionally, some participants tended to prefer repeating their literal or non-literal Finnish translations of the idiom instead of using the target language. All in all, most

repetitions, except figurative translations, were verbatim repetitions which accords with Lawson and Hogben's (1996: 120) findings.

Unlike other strategies, repetitions were not counted since this would have been extremely difficult as "true" repetitions would have been difficult to distinguish from reading aloud, talking aloud while typing or writing, and translation. For example, idioms were often repeated during translation, hence that other words in the sentence were translated but the idiom was left intact. Moreover, in reporting the final choice for the meaning of the idiom, some used the formula "idiom" is/means X, which technically is not repetition but may be difficult to distinguish from it. When the data was glanced for "actual" repetitions, it is evident that participants A, E and F used repetitions most. The TA protocols of A and E point to the use of repetitions especially prior to and in tandem with guessing the meaning. This is illustrated with the following excerpt from participant E's TA:

> (1) […] *sticka ut hakan*, vois olla johonki virheitten tekemiseen liittyvä, se ei oo, niin, se ei oo peloissaan, vaikka kuinka sais huonoa palautetta jälkeenpäin. *sticka ut hakan*, *sticka ut*, kuulostaa niinku vähän semmoselta, mikä se on niinku, öö ampua ohi. (Participant E)

This quotation illustrates that repetition can be used as a means to process the meaning of the idiom. It may also be used for trying to keep the idiom in memory. Nevertheless, there is a possibility that instances of this strategy were a product of using the TA method since the participants who produced richer and more natural TA seemed to be using this strategy more. Contrary to the positive evidence of the benefits of repetition, research (e.g. Gu & Johnson 2003; Fan 2003; Mizumoto 2010) consistently point to the negative relationship between visual repetition and vocabulary size and language proficiency. However, once again these findings may be more relevant for memorization than for meaning discovery. Still, repetition may play a minor role in meaning discovery but yet need to be accompanied with the use of other more efficient strategies.

8.4 Translation

Translation was considerably used as an encompassing and dynamic strategy for both English and Swedish. This is consistent with the findings of Pavicic-Takac (2008: 134.145) which showed that translation was an extensively used VLS for both German and English. Translation occurred throughout the process of discovering the meaning of an idiom. Instances of translation were not counted due to its ongoing and incremental nature. First, initial

translations were made of the words that were known. Then, the participant usually made guesses or looked up information in a dictionary. After this, the initial translation was revised by adding to or adjusting the previous translation. This procedure is demonstrated with three excerpts of participant G's translations of the idiom *cut no ice* and the example sentence (*Statistics cut no ice with anyone scared of going up in the air in a plane*):

> (2) "Tilastot eivät riko jäätä [unintelligible]. pelkää. ilmaa lentokoneessa." […] "Tilastot eivät riko jäätä, kenenkään kanssa joka pelkää mennä ylös ilmaan lentokoneessa." […] "Tilastoilla ei ole vaikutusta, eli tilastoilla ei ole mitään vaikutusta siihen, että, tilastot ei vaikuta siihen että paljonko ihmiset pelkää lentämistä." (Participant G)

In between the translations the participant completed look-ups in a bilingual dictionary, which explains the improvements made. Also note the advancement from literal to figurative meaning, a pattern found on a number of occasions. Because of the nature of the translation process described above, it would have been difficult to distinguish separate translations from each other. Even if the use of this strategy for a single idiom could have been recorded, other problems related to quantifying include its overlap with multiple other strategies such as repetition, guessing from context, and dictionary use. Additionally, the use of translation seemed to enable the use of other strategies. Firstly, the use translation enabled making use of the example sentence when not all words were known, hence enabling guessing from context. Secondly, translations were used as a means for testing the suitability of a particular meaning into the sentence, since many of the participants preferred to do this in Finnish instead of the target language. Thirdly, after translating the idiom into Finnish, and thus, having access to the literal meaning of the idiom, figurative strategies could also be used.

Various pieces of information were translated into Finnish. Most commonly, the example sentence given on a card, and the literal meaning of the idiom were translated into Finnish. However, multiple information abstracted from monolingual (English, Swedish) dictionaries were also translated, including definitions, synonyms, and example sentences among others. A review of the data reveals that almost all participants, particularly A, E, G, and I, used translation for the majority of both English and Swedish idioms. There do not seem to be major language-related differences, but rather differences to which extent individuals utilize this strategy. Furthermore, there is very little indication of participant D using translation. This may be due to a scarcity of TA data, and preference for quick dictionary lookups.

In sum, Finnish seemed to have a prevalent role in idiom meaning discovery. The high use of the translation strategy may be attributed to multiple factors such as language

proficiency, using Finnish for thinking aloud, perceiving meaning through L1, and defining the idioms in Finnish. It could also be taken as an indication of that translation, as a trace of grammar-translation method is still commonly used in Finnish upper secondary schools and language materials. However, this is only one explanation, and the findings of Pavicic-Takac (2008: 105-134) do not support this kind of a link between teaching and VLS.

8.5 Guessing

Overall, the second most used strategy was guessing meaning. It was mostly used for the target idioms, but also occasionally for single words. The guesses derived from several varying sources of information, such as lexical makeup of a word, the example sentence, or dictionary information including e.g. definitions and example sentences. Guessing seemed to have two differing functions: either discovering the meaning or finding a suitable formulation of the meaning. The latter function was likely triggered by the task, hence that some participants listed several possible meanings and then selected the most suitable one. This may also have been an attempt to understand the idiom more precisely. All guesses were uttered in Finnish, which all over again seems to be the language through which meaning was processed.

The guesses varied in terms of precision and confidence. Some guesses were vague, which was indicated by the use of "*is related to"* or *"has to do with"*. Other guesses aimed at giving the precise meaning of a word or an mwu. In figure 2 below, the number of guesses is presented:

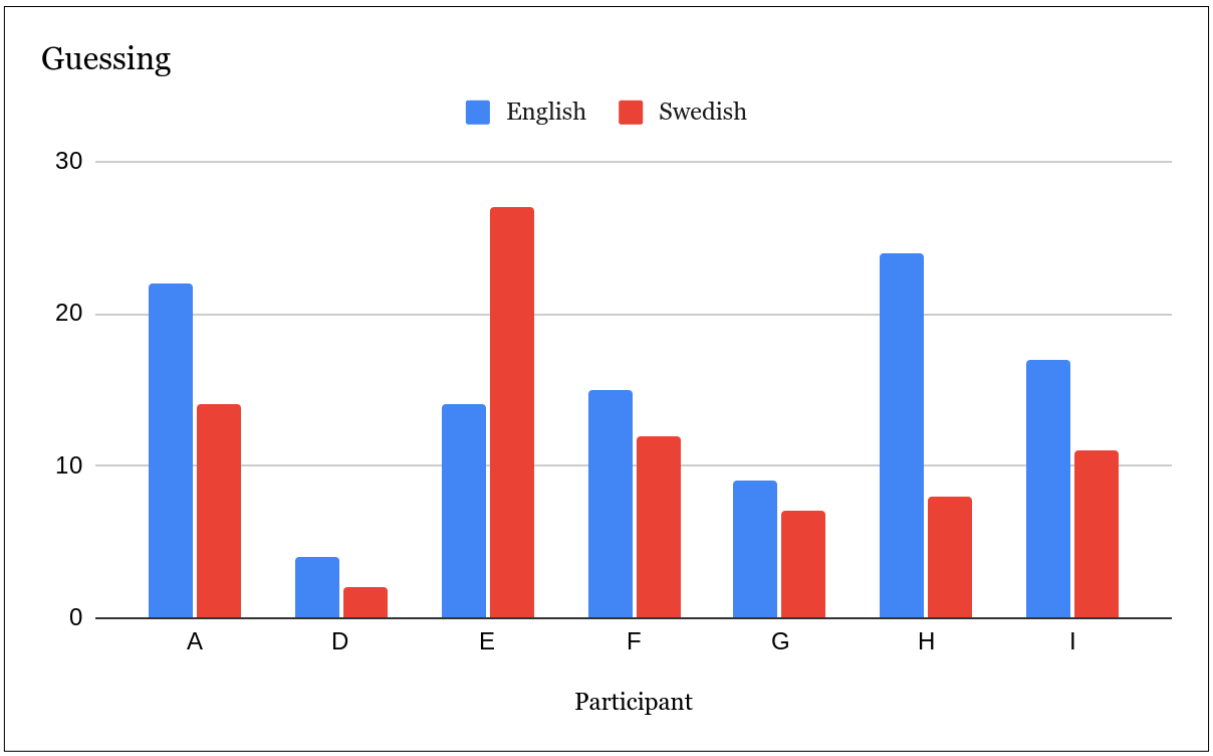

Figure 2. Guessing

As can be seen from the figure, overall, all participants excluding participant E, used more guessing for English than for Swedish. This is likely a consequence of understanding the example sentence better, which allowed more guessing from context to occur. Since the guesses were often affected by using a dictionary, the instances of the guessing strategy were further divided into two categories: before and after dictionary use. However, it must be pointed out that this does not mean that guessing from context never occurred after the idiom had been looked up in a dictionary. The number of guessing before a target word or mwu was looked up in a dictionary are shown below in figure 3:

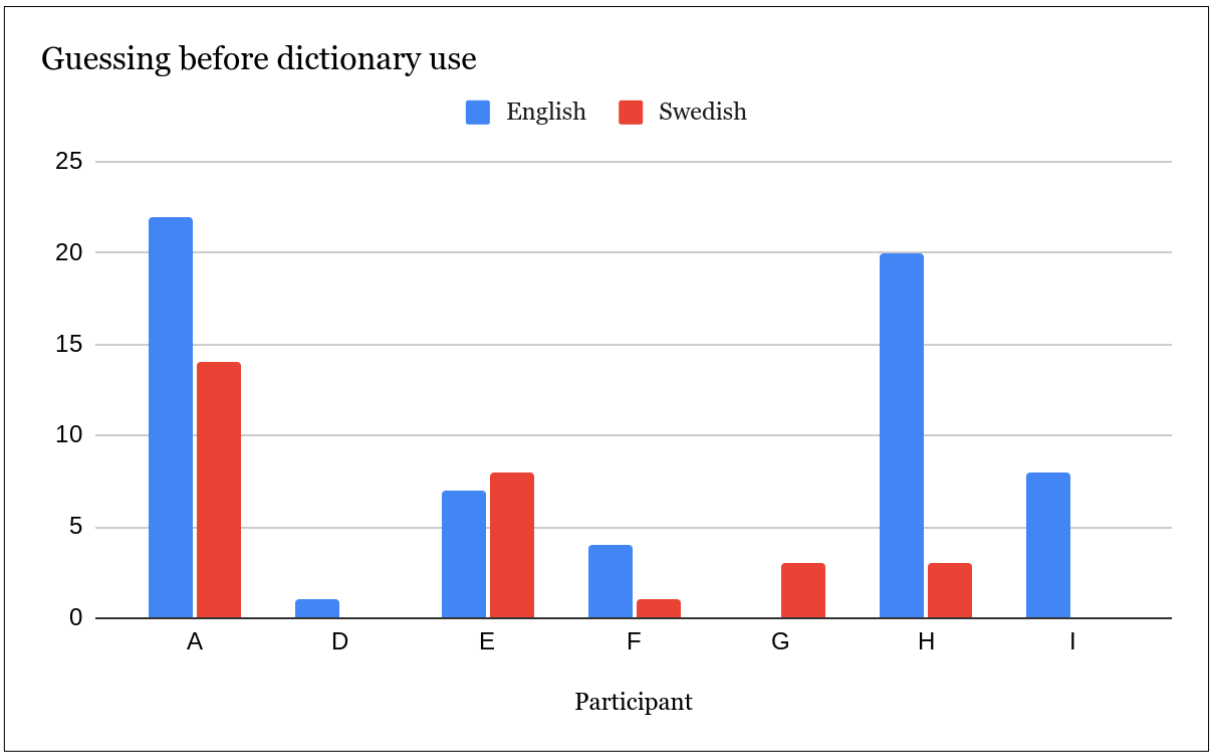

Figure 3. Guessing before dictionary use

The figure reveals a somewhat different view compared to the figure 2. For instance, for participants H and I, the gap between languages grows. In general, the participants seemed to guess the meaning more likely for English when they had not yet consulted a dictionary. What is even more interesting is that there were vast differences between individuals. For example, participant A's main strategy was guessing from context, and he did not consult a dictionary even once neither for English nor for Swedish idioms.

Guessing before searching for answers in other sources was made possible through utilizing a variety of clues. Idiom guesses were often based on the example sentence, the literal meaning of the idiom, or background information. Quick guesses of single words right before dictionary consultation also occurred at this stage. These guesses could have been either guesses of unfamiliar words or retaining the meaning of a word the participant was insecure about. After dictionary consultation, usually the rejection or confirmation of the guess was somehow commented on. When examining the number of guesses made after dictionary use, the results even out. This is shown with figure 4 below:

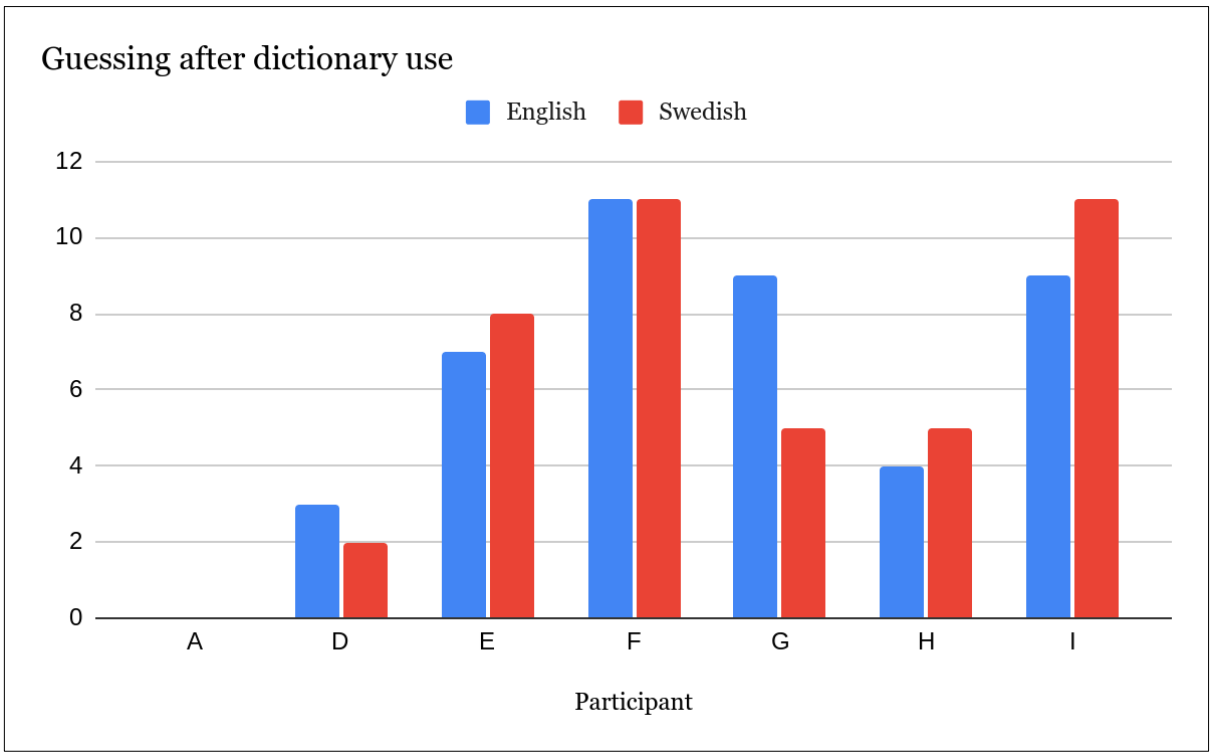

Figure 4. Guessing after dictionary use

Surprisingly, guesses of Swedish idioms were even more common for participants E, H and I. These guesses were often made later in the meaning discovery process. They often seemed to be attempts at finding a suitable Finnish word or phrase for the idiom. Participant E was particularly keen on this, and when she was not pleased with some of her guesses, new ones were formulated till one proved to be good enough. Whereas others began the meaning discovery with forming one or several hypotheses of the meaning, others preferred to gather data from dictionaries and other sources prior to forming their hypothesis.

A wide range of clues were used to infer the meaning of the idiom. For some clues there was explicit evidence in the data, as revealed by the use of some logical connectors such as "because". Other clues were found by "reading in between the lines", that is examining what was uttered or done before and after the guess. Five types of clues, or information were found from the data: (1) using example sentences, (2) using background information, (3) using literal parts of the idiom, (4) using metaphor or origin of the idiom, and (5) using dictionary information. At times, multiple clues were used in tandem to make a guess or a sequence of guesses. Moreover, participants made comments about how "logical" or "sensible" something is/seems/sounds. This could be also considered as one type of clue, using word knowledge.

Using the example sentence from the card or provided by a dictionary involved guessing the meaning from context. It was the most used clue for guessing. Whereas some

appeared to have used the whole sentence, others seemed to identify words or parts of the sentences as relevant clues. The example sentence (When the police arrived, he took it on the chin apologizing for the trouble he had caused them) for the idiom, *take it on the chin*, strongly directed the participants' guesses. Key words mentioned by some of the participants were *police* and *apologize*, which were well reflected in the following guesses: "take responsibility", "start explaining", "admit one's mistake", "humble oneself", and "consent". Background information, i.e. word knowledge, of an encounter with law enforcement may also have been used. Couple other more evident examples of using background information were also found. In one of these examples, participant A pondered the meaning of the Swedish word for garden by testing whether the literal meaning of the idiom *ta itu med något* would make any sense in the sentence. His reasoning and inferencing is illustrated with the following quotation:

> (3) "*Ta itu med*, mikäs tuo trädgården on, voiko sitä ottaa paloiks?" […] "Nyt minun täytyy ottaa palasiksi puutarhan kanssa. öö, vai onks se niinku puutarhassa yleensä tehdään jotain joten se vois olla ryhtyä tekemään jotain." (Participant A)

The quotations demonstrate that the participant ruled out the literal translation with the help of background knowledge of gardening. Alternatively, it is possible that the participant may have had some earlier information of the idiom *ta itu med något*, which is quite common in Swedish, since he abruptly arrives at the correct meaning without using dictionary. Other instances of using background information involve associating the meaning of the idiom with a certain subject. For instance, participant E reads the definition of *give someone a run for the money* in a monolingual English dictionary but instead of trying to understand it, misleadingly interprets that the idiom has something to do with "betting". Albeit this misunderstanding could be attributed to other reasons such as lack of vocabulary or literal interpretation, it is very likely that the participant associates the idiom with the word "game" that occurs in the dictionary definition. Then, by using background knowledge of "money" and "game" the participant arrives at the erroneous interpretation, betting.

Almost all participants used the literal meaning of the idiom to guess its actual meaning. The use of this strategy has also been observed in two similar studies on idioms (Cooper 1999; Katsarou 2010). After having figured out the literal translation of *sticka ut hakan*, participant F proceeded from first, stating the translation "stick your chin out", to then making the guesses, "come forth", and "express one's opinion". These guesses were additionally justified with a reference to the cause-and-effect function in the sentence, which holds that 'sticking your neck out' may result in receiving criticism. Participant F was clearly on right track with his guess, whereas at other times, the connections made between the literal and the figurative meaning do

not make sense at all. As an example of this, participant A firstly, drew a conclusion that lock, stock and barrel meant "immediately directly". However, he was clearly not satisfied with this guess, and consequently re-read the sentence and made further guesses. Then, he drew another parallel between the literal and the figurative meaning by repeating his first guess with a slightly different, and peculiar wording: "locking into a stock and a barrel" and "quitting at a time to those places", which supposedly means leaving things as they were. The use of this strategy did not clearly work as well with more opaque idioms, whose literal connection has been lost (cf. Grant & Bauer 2004).

Only two instances of using metaphor as a clue for guessing were found, both uses by participant H. In these uses, he richly used the imagery and described the metaphor with varying degrees of success. Firstly, regarding the idiom *make a splash*, he described a view of someone jumping into a pool and making a huge splash. Based on this, and likely the example sentence, he made several guesses: "awake interest or attention", "a thing that is significant or famous", and something that "makes a big impression". Although he was on right track, lack of dictionary use nullified the promising start set by the guesses. Secondly, participant H made an interesting interpretation of the metaphor in the idiom *bend over backwards*. He associated it with thanking and bowing as when someone "went double" to show politeness. Despite the meaning being incorrect albeit fitting perfectly to the context, and the hypothesized origin being wrong as well, the participant showed understanding of the fact that the origin of the idiom can be metaphorical.

In sum, the participants used various clues to guess the meaning of single words and mostly target idioms with varying success. Idiom specific strategies were, nevertheless, used less than, for instance, using the example sentence. Overall, guessing seemed to be a popular strategy, particularly for English idioms. However, as illustrated with examples from the data, guesses often lead to overgeneralizations of the contextual meaning (see also Szczepaniak 2006). Additionally, some initial guesses were sustained despite contradictory evidence found in other sources (cf. Macis 2018).

8.6 Using a search engine (googling)

Using a search engine was used as a pre-strategy for dictionary use. All the instances of using this strategy, except for two searches with a different search engine, *Duckduck.go*, included using Google, which is why this strategy will henceforth be called googling. Googling is

separated from dictionary use, since using it involved making different strategic decisions, such as deciding what accompanying search terms will be used. The results of the searches also seemed to play a decisive role in what dictionaries were made use of, since most dictionaries were accessed as first results of a google search. If search terms, such as "meaning", are not considered, nearly all searches included more than two words, most commonly the full idiom, or parts of the idiom. One exception to this is participant H who googled four single words. This finding indicates that the participants approached the idioms holistically viewing them as multiword units. The typographical enhancements in the word card may have promoted this (see also Bishop 2004). Moreover, the participants may have tried their luck and hoped for quick answers for the whole idiom. The number of google searches are shown below in figure 5:

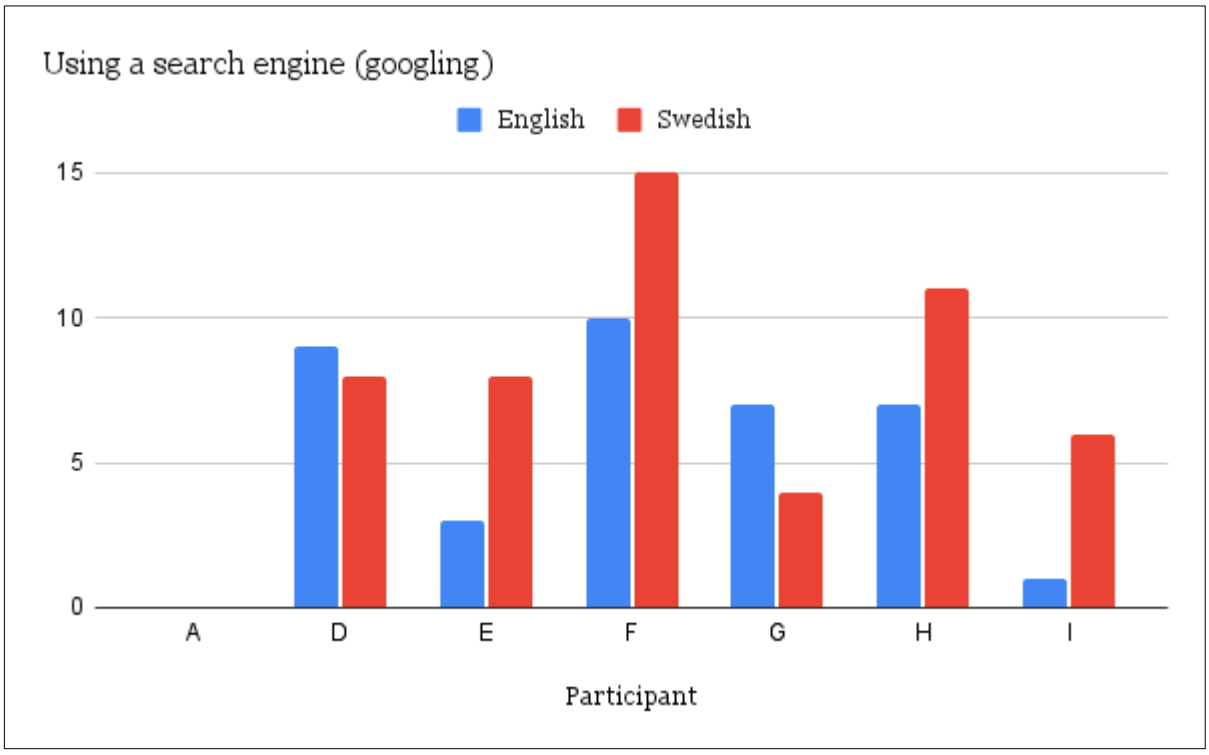

Figure 5. Using a search engine (googling)

Six of the seven participants utilized Google search as a strategy. As can be seen from the table, there were both individual and language differences. However, no clear patterns could be detected. Firstly, individuals seemed to differ in their preferences of this strategy. One possible explanation for this may be differences in approaches. Some may prefer to complete a new Google search, whereas others may prefer turning to already open websites and dictionaries for help. Secondly, the majority of the participants (E, F, H and I) used googling more for Swedish

idioms with a slightly higher use compared to English, whereas two of the participants (D, G) used googling more for English idioms.

Unlike the other participants, who opted for using Google, the default search engine, participant F demonstrated that he was also able to use a different search engine, duckduck.go as well. This change shows that he was able to evaluate the strategies used and make adjustments in order to test if a new strategy would prove to be more helpful. Nevertheless, changing the search engine did not seem to bring relief to his task performance.

In googling, the participants portrayed a use of a variety of accompanying search terms. These are listed in table 5 below:

Table 5. The number and type of search terms used when using a search engine

Participants	English	amount	Swedish	amount
D	meaning in finnish [*sic*]	6 3	translation på finska engelska på engelska	2 4 1 1
E	-	-	betydelse engelska	1 1
F	suomeksi meaning	2 2	fraaseja suomeksi “” “” suomeksi	1 4 2 2
G	meaning tarkoittaa	2 1	merkitys betyder vad betyder	1 1 1
H	suomi suomeksi	3 3	suomi suomeksi synonym	2 6 1
I	meaning	1	på engelska meaning betyder suomeksi	1 1 1 1

Table 5 shows that participants differed in their use or selection of search terms. Overall, the participants searched for Finnish translations and meaning, which is not a surprise in view of

the meaning discovery task and the choice to define the idioms in Finnish. Additionally, participants D, E and I searched for English translations, which will be discussed as a separate strategy in section (8.10). The choice of search terms is significant since they play an important role in the search results. For example, the search terms *meaning* or *betydelse* 'meaning' often resulted in monolingual results, whereas the use of *suomeksi* 'in Finnish' likely led to bilingual results. One remarkable finding that needs to be pointed out was participant F's ability to use quotation marks. It is not only a sophisticated searching strategy, but also an indication of that the words in the idiom were sought in sequence, not separately.

But how did the participants end up using these search terms? Two alternative routes for this emerged from the VSC recording. Either the participants used search terms that they were familiar with, or they used incremental search as a search technique. The latter involves "automated term completion from an index of available terms, before the complete term is typed" (Lew 2013: 24). In other words, the suggestions made by Google were used. Below in Figure 6, a screenshot of one of these indexes is shown, to illustrate the wide range of options participants had to select from:

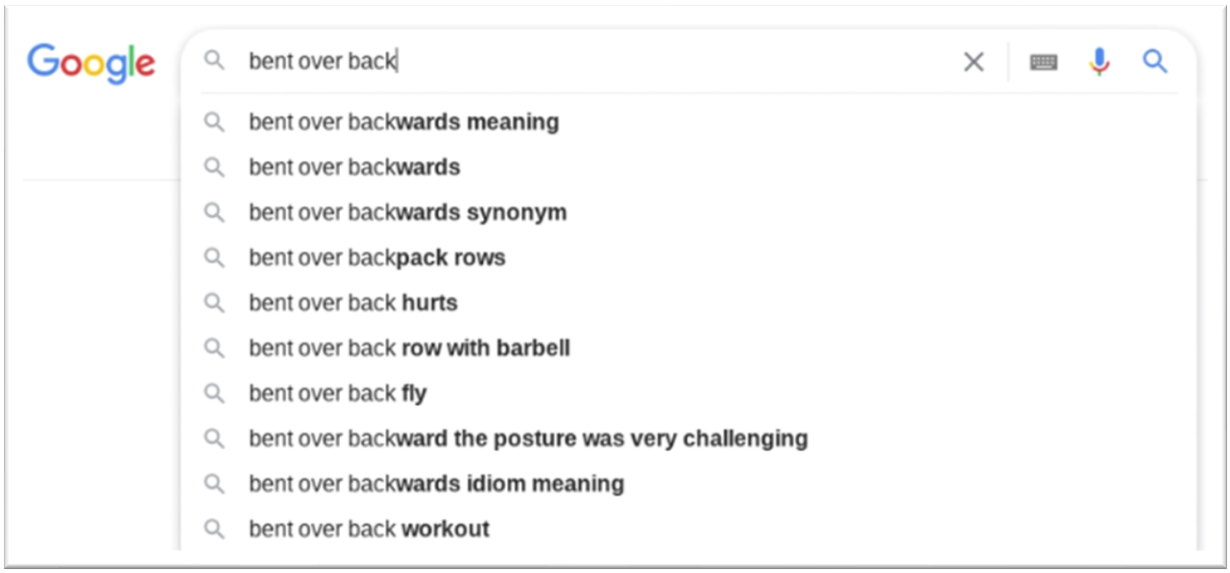

Figure 6. Google search indexes for the idiom "bent over backwards"

The screenshot also reveals a couple more interesting points. Firstly, the second last alternative indicates that the phrase is an idiom, a piece of information that could have been very useful in completing the task. Secondly, the participant used the inflected form of the verb, i.e. idiom variation, in the search. Multiple other variations or erroneous forms of the idioms and single

words were typed both when googling and using dictionaries. With regards to idioms, the following types of errors were found: spelling error or copying error (*göra en djuådykning*), idiom variation from the example sentence (*bent over backwards)*, and own erroneous idiom variation (*jump a gun*). Similarly, when searching single words, inflected forms of the word were frequently searched. In some cases, these errors resulted in failure of the search, but even more often, the fuzzy-spelling search feature (Lew 2013: 26) offered corrections and suggestions. These turned out to be a significant aid for the participants, since these features nullified the gravity of the errors made.

8.7 Monolingual dictionary

Monolingual dictionaries were initially always accessed via Google. After the dictionary entry had been opened, most participants closed it or proceeded to complete next searches with Google, whereas participants E and I chose to complete later idiom look-ups with the previously opened Cambridge Dictionary. Furthermore, as mostly mwus were googled, all monolingual entries displayed idioms. In other words, meanings of single words were not looked up in monolingual dictionaries, solely idioms. The number of monolingual dictionary look-ups are presented in figure 7 below:

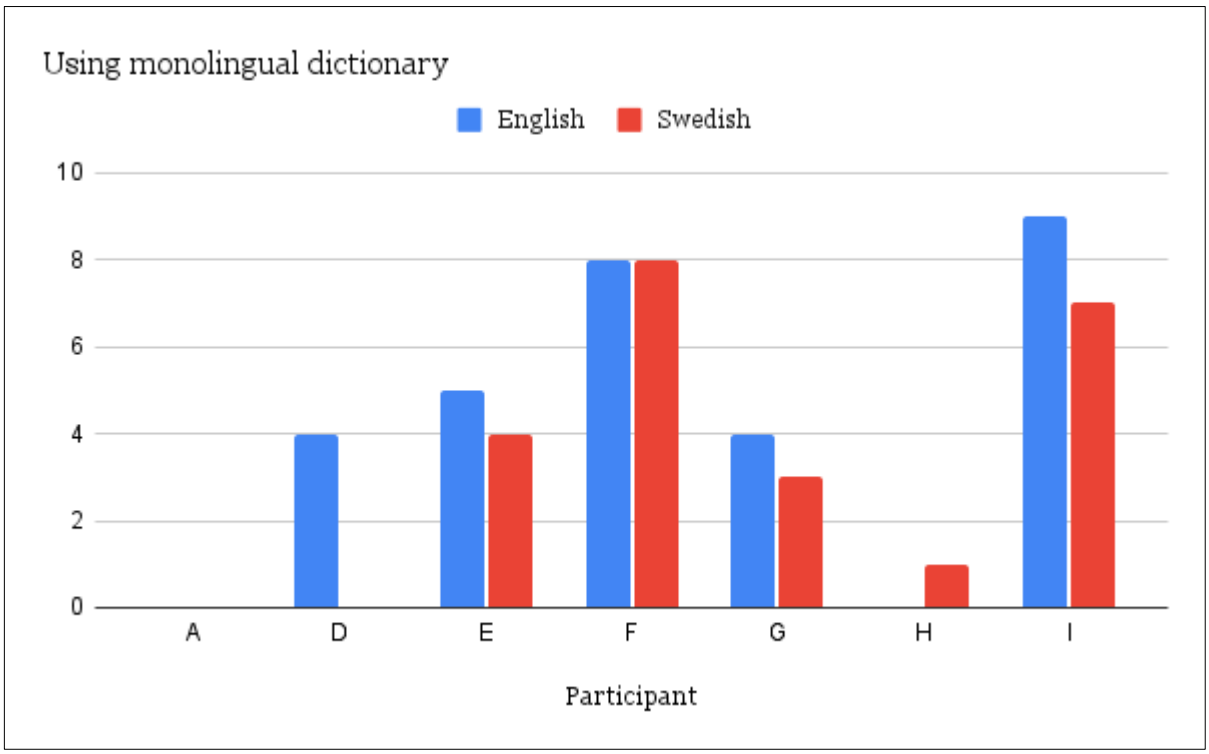

Figure 7. Using a monolingual dictionary

All participants, except participant A who only used a bilingualized dictionary, at some point used information from a monolingual dictionary. All in all, monolingual dictionaries were more used for English than for Swedish idioms. Participant D for instance, did not use monolingual dictionary at all for Swedish. An explanation to this, as suggested by earlier research (Atkins & Varantola 1998a: 43; 1998b: 113; Lew 2004), could be participants' higher proficiency in English.

Various conventional dictionaries and less prestigious dictionary sites (see Nesi 2012 and Yongwei 2012) were used to discover the meaning of both single words and mwus. The high variability in the types of dictionaries can be explained with accessing entries via Google, since the search results can vary each time. Some of the English dictionaries were prestigious dictionaries such as *Cambridge Dictionary*, and *Merriam Webster Dictionary*. The monolingual dictionary which was used the most was Google dictionary, which cooperates with *Oxford Languages*. The reason for its high use may lie in its prominence in Google search as the first search result. Figure 8 below shows an example of a Google dictionary entry for the idiom *bend over backwards*:

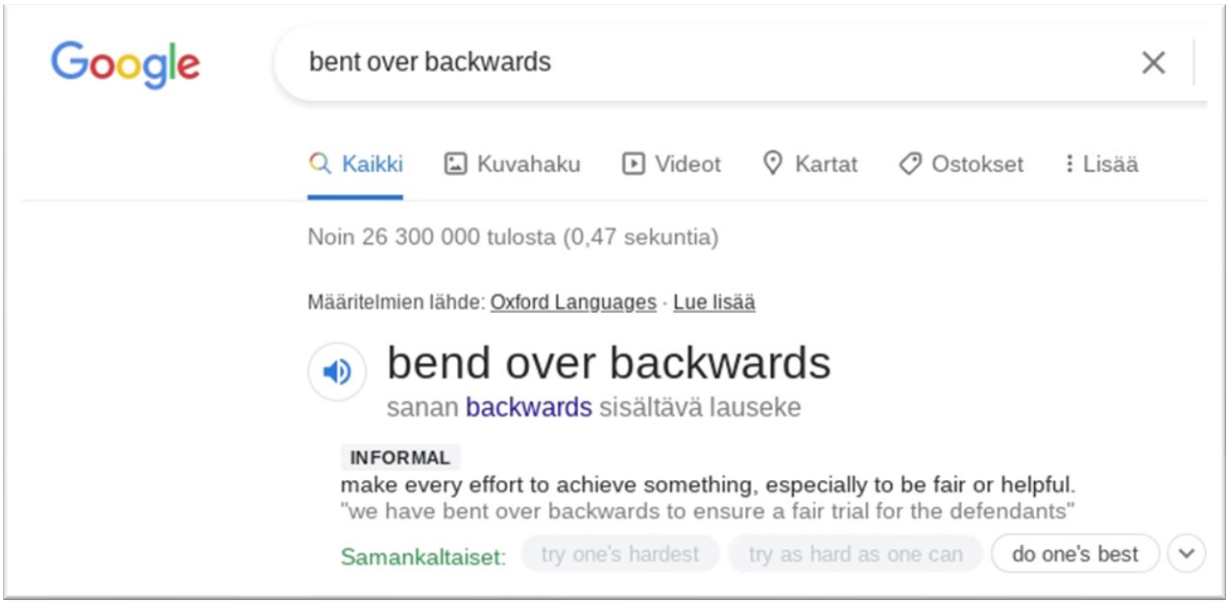

Figure 8. Google dictionary entry for the idiom "bent over backwards"

From these search results, the dictionary user can already find a definition, information about formality, an example sentence, and a synonym for the idiom. The easy accessibility to required information probably explained why not many participants actually clicked on the suggested websites but rather read dictionary entries directly from the search results. Another English monolingual dictionary that was used was *Cambridge Dictionary*, which was the main source of information for participant I. Other English dictionaries were non-lexicographical and non-academic online dictionaries, such as *Dictionary.com*. The quality of Swedish monolingual dictionaries seemed to be somewhat lower. The dictionaries that were used inluded several uses of the freely edited wiki, *Wiktionary*, two synonym dictionaries (*synonymer.se*, *typkanske.se*), and a monolingual dictionary (*ordguru.se*) among others. No idiom dictionaries, such as the alternative e-dictionary (Nesi 2012), *Free Idiom Dictionary* were opened.

All participants used various information from monolingual, bilingual, and bilingualized dictionaries. The TA revealed that some participants were clearly conscious of the various information provided by a dictionary in that they intentionally looked for translations, synonyms and/or example sentences. As for monolingual dictionaries, definitions were used, that is read aloud the most. Example sentences, when simple enough, provided additional material for guessing from context. One successful use of example sentences was illustrated by participant I. She read three example sentences for the idiom *en annan femma* in the Swedish *Wiktionary*. Two out of three example sentences proved to be helpful enough to guide her to guessing the correct meaning of the idiom, alongside using translation. Etymology

of both English and Swedish idioms was read aloud altogether five times but in none of the instances the participant further commented or elaborated on this information. This was taken as an indication that the information was probably not understood.

Consulting a monolingual dictionary was not unproblematic. On multiple occasions, the participants were not able to make full use of the information provided by the dictionary. For example, participant E opened a *Wiktionary* entry for the Swedish idiom *sticka ut hakan* and straight jumped into read the end of the first example sentence. After this, she proceeded to read the beginning of the sentence. Next, she read part of the etymology, and then made a comment that she understood nothing of it. Her careless reading led her to dismiss two instances of the definition, and some synonyms that may have been useful. Thus, selective reading was one cause for problems and it occurred in both languages. Another problem experienced by the participants with both Swedish and English idioms relate to comprehension of the entries. If an entry was not comprehended, it was usually either dismissed, or unknown words were translated. For example, participant F was not familiar with the synonyms *förkovran* and *inträngа* provided for the Swedish idiom *göra en djupdykning* and looked up both words in a bilingualized dictionary with little success. The data shows that using monolingual dictionaries took time and energy. Moreover, if an entry is not fully understood determining the reliability of the dictionary, and relevance of the entry may also be difficult. Nevertheless, it can be assumed that more information of the idioms was available in the target language than in Finnish, which is why using a monolingual dictionary occasionally proved to be helpful.

8.8 Bilingual or bilingualized dictionary

Bilingual/ized dictionaries contained information always in the L1, or both in L1 and L2, i.e. English or Swedish. This included the use of Google translate, since apart from allowing searches of longer phrases and sentences, it provided either one or multiple translations to the language chosen. Bilingual and bilingualized dictionaries were used the most of all the strategies. The number of dictionary uses are presented in figure 9 below:

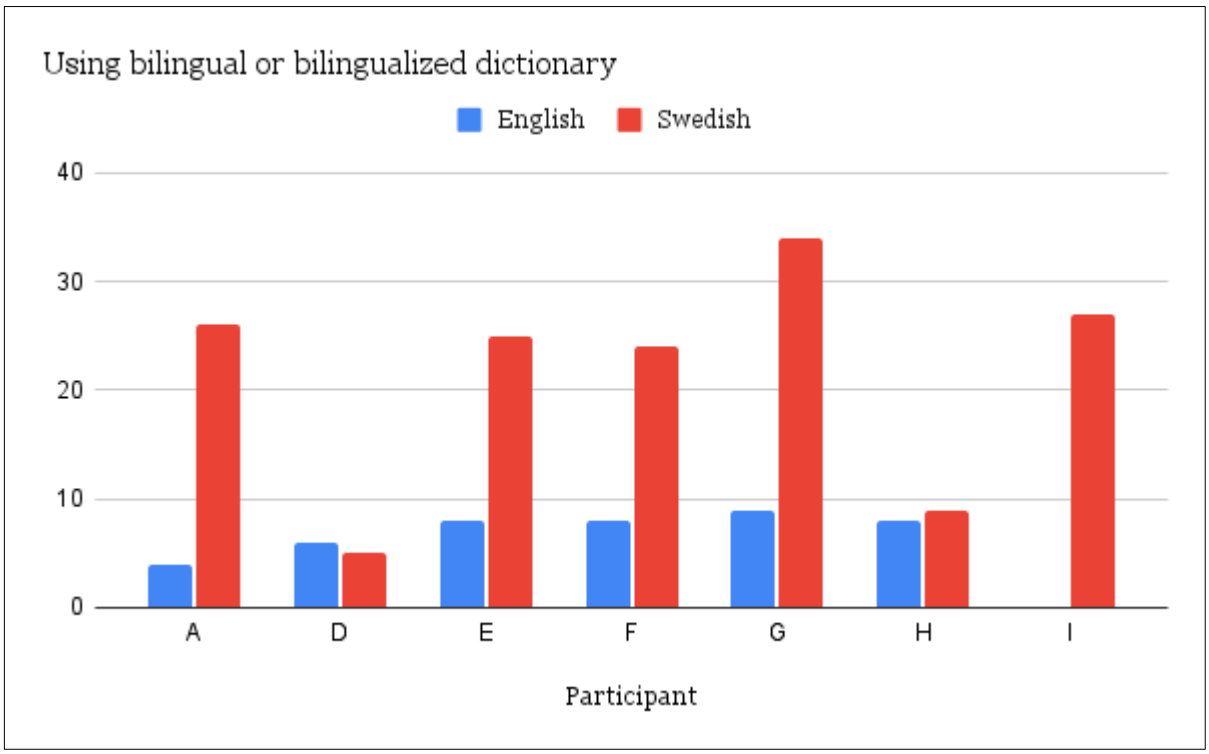

Figure 9. Using bilingual or bilingualized dictionary

From figure 9 above, it can be seen that bilingual/-ized dictionaries were used significantly more for Swedish than for English. In general, a very high number of instances of this strategy were found from the data; five participants used this strategy fairly over 20 times. This means that for these participants at least four dictionary searches were recorded per one Swedish idiom. However, not all participants used biligual/ized dictioanries as extensively, or as their main strategy. Participant H mostly resorted to guessing strategies, and participant D was content with one *Google translate* look-up per idiom, regardless of the quality of the translation. Others appeared to be fixated on translating every unfamiliar word, which inevitably resulted in multiple dictionary look-ups particularly for Swedish.

Bilingual dictionaries were consulted for both single words and mwus, including novel, creatively constructed mwus (*ner foten*) and institutionalized mwus (*hold out*). These lexical items were most commonly taken from the target idiom, the example sentence, a dictionary or a webpage. It can be hypothesized that the dictionaries were mostly used to discover the meaning of unknown lexical items, meanwhile the high number of dictionary uses seem to hint on other possible reasons as well. For instance, it is possible that the participants, of whom some were aware of the fact that words can be polysemous, consulted dictionaries multiple times if they were not satisfied with the information obtained, and looked for other meanings that would better fit in context. Additionally, some items may have been forgotten. Guesses of

meaning before dictionary consultation showed that not all items were unknown but confirmation for existing information was also sought. Single words were sought substantially more than mwus or idioms. Additionally, the participants clearly preferred to look up the meanings of single words in the bilingualized dictionary *sanakirja.org* and the bilingual translator, *Google translator*. In contrast, idioms and other mwus were additionally looked up in several other dictionaries, which appeared in Google's search results. Hence, a wider selection of dictionaries were used for idioms.

Several different bilingual/ized dictionaries were used but *sanakirja.org* and *Google translate*, were used by far the most. The former was used 92 times, and the latter 74 times of altogether 193 uses of bilingual/ized dictionary, which accounts for roughly 86% of all uses. The first most used dictionary was the bilingualized online dictionary *Sanakirja.org* that provides the user with various information, such as translations, definitions, example sentences, and grammatical information, depending on the item. The second most used dictionary, or rather bilingual translator was *Google translate*. One reason for its high use may be its support for mwus. Interestingly, it was extensively used particularly by D, G and I, despite some of them questioning the quality of the translations. However, a crucial difference was observed: Participants D only used Google translate for Swedish idioms, whereas participants G and I utilized parallelly other dictionaries and websites for comparisons. Other bilingualized dictionaries that were used included the bilingualized online dictionaries *Ilmainen sanakirja*, *bab-la*, *Redfox*, bilingual context dictionary *tr-ex*, and bilingual translator *Linguee.com*. At a first glance, all of these bilingual/ized dictionaries and translators, except *Redfox*, seem to be non-academic and non-lexicographical online dictionaries. This finding could explain some of the erroneous translations these dictionaries provided. But at the same time, they were accessible, free, and at times helpful in meaning discovery (see also Moon 2015).

Even if bilingual/ized dictionaries were multiple times successfully used, many problems were identified from the VSC recording. This is unsurprising, since lack in dictionary skills has also been observed in other studies (e.g. Wingate 2004), and previous research has shown that mwus and idioms pose a challenge for dictionary users (e.g. Chen 2016; Li & Xu 2015). Firstly, as reported earlier, idiom variations and inflected forms of words were looked up with varying success. Secondly, wrong senses for single words were chosen (see e.g. Chen 2016). As an example, participant A assumed that the word limb, in *go out on a limb*, meant an appendage, not a branch of a tree, a conclusion to which he arrived after checking both translations and English definitions for the word. The prominence of the wrong sense as being

more familiar and being the first translation, may have led participant A to the wrong choice. Additionally, as will be discussed later, participants had a tendency of choosing senses for the constituent words of the idiom based on their literal suitability in the example sentence, a strategy which was of little use since idioms are non-literal expressions. Thirdly, at times the participants were unable to evaluate the appropriateness or quality of information found. For example, two participants accepted the incorrect translation *kierteessä* 'in a spiral' of the idiom *in the loop* that was provided by Google. Fourthly, some information, that may have proved to be the key for correct meaning was left unattended or ignored as irrelevant (cf. Szczepaniak 2006). Lastly, the enormous number of bilingual/ized dictionary uses suggest that some of the consultations were irrelevant (cf. Prichard 2008).

8.9 Using other websites

Four of the participants also used other websites for various purposes altogether nine times, which was very minimal compared to dictionary use. These non-academic and non-lexicographical webpages were three encyclopedias: *Wikipedia, Nissepedia*, and *Encyclopedia Titannica*, two chat forums: *Suomi24*, and *ProZ.com*, a personal webpage and *Quizlet*. The usefulness of these websites varied a lot. For instance, participant E tried to understand *Nissepedias* monolingual and wordy explanation of the idiom *göra rätt för sig* by translating unknown individual words occurring in the definition, yet with little success. Furthermore, participant H explored two chat forums, which provided possible meanings for the idiom *lock, stock and barrel*. Below in figure 10, Finnish translation suggestions for the idiom made in the *suomi24* forum are displayed:

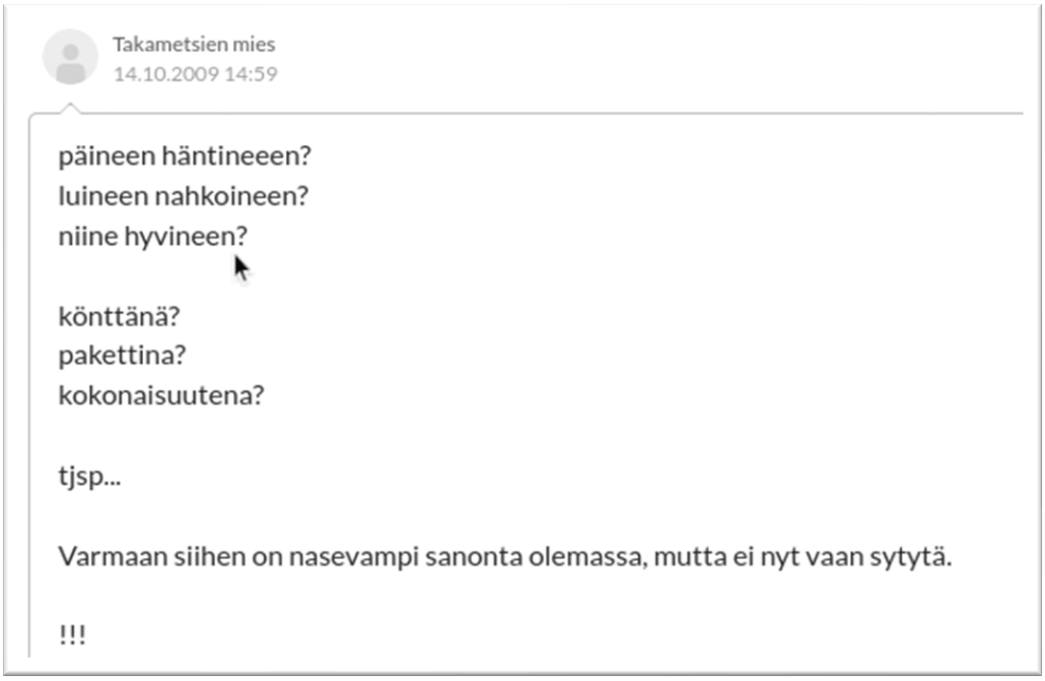

Figure 10. Suomi24 translation suggestions for the idiom "lock, stock, and barrel"

These careful suggestions made by an anonymous person led the participant to guess the meaning of the idiom correctly. This observation confirms Moon's (2015) analysis that informal websites can provide sufficient information about idioms despite some of their lacks in quality.

In contrast, other websites, often non-related to idioms, were of little help. For instance, participant F very quickly abandoned the Quizlet page as irrelevant, since it only contained flashcards of single words, and participant E did not even click on a personal webpage titled *Festina Lente*, which is *skynda långsamt* in Latin. An interesting example of how searching the meaning of idioms may result in non-linguistic information, was participant D's search after the meaning of stoicism. It was triggered by not comprehending the word *stoically* occurring in Oxford's monolingual definition of *take it on the chin*. Eventually, it led the participant to read about stoicism in Finnish, again with not much help for the actual task. These results exemplify, on the one hand the flexibility in the sources used for discovering the meaning, and on the other hand the importance of skills in evaluating the cost and benefit ratio of using the webpage.

8.10 Using English

This strategy included using knowledge of English as a means to find out information about Swedish idioms. A similar strategy found in the VLS literature is Labontee's (2019) utilizing lexical knowledge of other languages. Using English was a strategy that was specific for Swedish. It manifested itself in Google searches and using dictionaries but also in reading information in English. Thus, it is a parallel strategy of googling, dictionary use, and using other websites. Four participants, D, E, H, and I, used this strategy. Of these, participants E and H used it only on one occasion, and participant D used it only six times when googling and using google translate. In contrast, the strategy was much more extensively and elaborately used by participant I. She used the strategy altogether 20 times, and it was one of her main strategies for figuring out the meaning of Swedish idioms. Almost as much as using *Google translate* from Swedish to Finnish, she parallelly utilized and compared translations from Swedish to English. If the translations matched, information was regarded reliable. However, when the Finnish and English translations considerably differed, participant I relied more on the English translation. Interestingly, for two Swedish idioms *dött lopp* and *se om sitt hus*, translations of English idioms, *dead heat* and *feather one's nest*, were provided. These triggered additional searches of meanings of these English idioms. In figure 11, the screenshot of the definition provided by Cambridge dictionary, which lead participant I to the correct meaning, is shown:

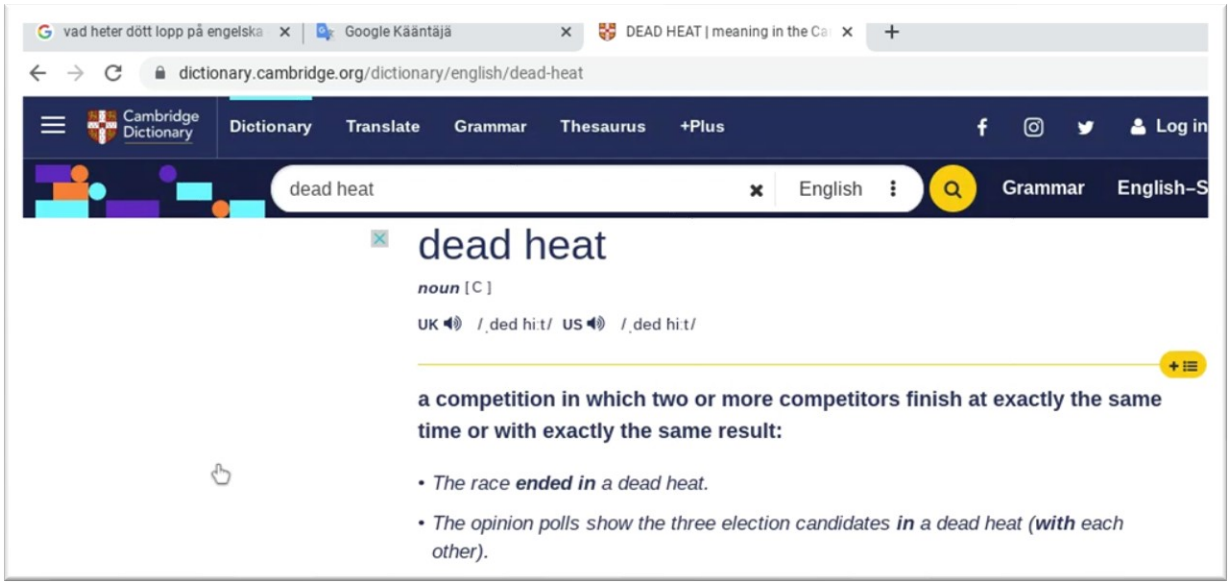

Figure 11. Cambridge Dictionary definition for the idiom "dead heat"

The participant was able to derive the meaning *tasapeli* 'tie' by reading the definition. Additionally, the initial Google search, *vad heter dött lopp på engelska* 'what is *dött lopp* in English', is visible in the left upper corner.

English is a world language. By the use of English, the participants showed that they were aware that more information is available online in English than in Finnish. For instance, participant I even comments that "English sayings can be easier found". Moreover, the use of this strategy indicates that the participants were competent in using three languages simultaneously, and it did not seem to be an additional burden to them. These findings, alongside other findings confirm that English proficiency was much higher.

8.11 Other strategies

In addition to reading aloud information about the etymology of the idiom, only one occasion of etymological elaboration was recorded. The strategy was used by participant H to confirm the Finnish translation of the idiom *jump the gun*. He points out that "even in sprinting or other races, when you go before the gun so that's a false start." Etymology or origin of the idiom can be more easily inferred after the meaning is known. Altogether six instances of retrieving single words in mwus or sequences were also recorded. This was mostly typical for participant E, who thought aloud the following phrases "scratched furniture", "*pappa betalar"*, "*jag var alltid rädd i skolan*", and "*upp med dig*". These phrases indicate that the single words relevant for the task, such as scratch, were originally learned in chunks. Additional examples of the use of this strategy were listing Swedish irregular verb tenses, such as *hålla, håller, höll, hållit*, which are often learned by heart. Furthermore, two instances of analyzing the components of a compound word were recorded.

8.12 Summary

Rather than separately employing single strategies, the strategies were used in sequences and clusters, hence overlapping with each other. These findings resemble the nature of language learning strategy use described in Cohen and Wang's (2013) study. Overall, the use of strategies involved forming and testing several hypotheses (see Cooper 1999). The formulated hypotheses were evaluated and compared with the information found. As a result of this, a new hypothesis was formed and the old one abandoned, or the old hypothesis was confirmed.

Moreover, the learners had a problem-solving approach to the task (see Cooper 1999). Strategic actions were used as a means to "crack the code of the idiom", and if a strategy did not seem helpful, a different strategy was tried instead. As idioms turned out to be difficult for the participants, solving the problem of "discovering their meaning" at times required creative solutions, and exploration of a variety of strategies, which is reflected in the large selection of strategies, and substrategies found in this study.

All participants engaged in the task differently, dynamically employing different types of strategies different number of times, which shows that strategic actions depend on the person (Gu 2003b). Participants A and H used guessing strategies extensively, however, participant A did not confirm his guesses with dictionaries as participant H and E did. The order of guessing and dictionary consultation also varied, for instance participants F and I preferred guessing the meaning of the idiom after firstly gaining some information online. There were also qualitative differences in the number of strategies and the variety of strategies employed. As an example, participant D's strategy repertoire was scant, as she mostly used googling, and some bi- and monolingual dictionaries. In contrast, participant E's strategy repertoire was far more comprehensive and richer, encompassing almost all of the strategies in the strategy categorization. Participant F could be named as a "computer genius", since he demonstrated the use of ingenious and sophisticated searching strategies and showed that he was able to critically evaluate webpages and dictionaries, a skill which the other participants also would have benefited from. Some strategies were more specific for certain individuals: participant E assessed her knowledge of words and mwus several times, participant G often resorted to suspending the meaning discovery of Swedish, and participant I mainly used English as a means to search information about Swedish idioms. Other strategies, such as using monolingual dictionary for English, and bilingual dictionary for Swedish, were shared by almost all participants. The use of strategies, and the language differences are summarized in figure 12 below:

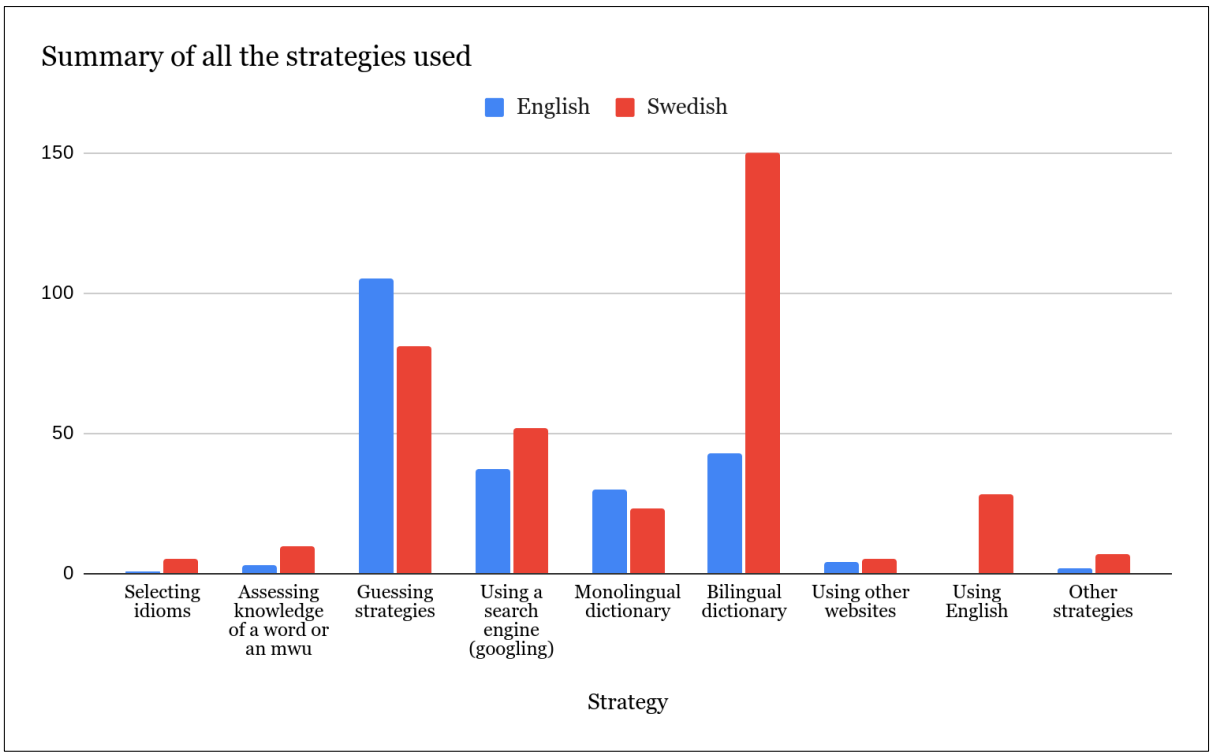

Figure 12. Summary of the strategies used

The data for the meaning discovery of altogether 48 English idioms, and 35.5 Swedish idioms suggest that VLS for idioms depend on the language. The strategies that were more used for English were: guessing strategies and using monolingual dictionary. All other strategies, particularly using bilingual dictionary, were more used for the meaning discovery of Swedish idioms. Furthermore, a strategy, entirely specific for Swedish emerged from the data, namely the use of English for searching information on Swedish idioms. All in all, top four strategies were (1) bilingual dictionary, (2) guessing strategies, (3) using a search engine, and (4) monolingual dictionary. Other strategies were much less used.

9. TASK SCORES AND ERROR CATEGORIES

In this section, the answers of the meaning determination task are analyzed. First, task scores will be presented. After this, the error categories for partially or fully incorrect answers will be described and illustrated with examples. Throughout this section, differences between English and Swedish will be pointed out and discussed alongside a comparison of individual learners.

9.1 Task scores

All participants chose to give define the idioms in Finnish, although they were allowed to use English and Swedish as well. This may be due to their heavy reliance on L1, as shown by the use of translation strategy. Another explanation may be found in the instructions. The participants were instructed to move on to the next idiom when they felt that they had understood the idiom. This implies that the participants may have aimed at understanding, which necessarily may not have been reached if monolingual definitions had been copied to the letter. The task of defining the meaning of the idiom elicited multiple suggestions or guesses before a final conclusion was reached. When inspecting the TA protocols closer, it has to be noted that some participants clearly had partial or full understanding of the meaning even if it is not reflected in their answers. Furthermore, some of the guesses were correct but they were for some reason later abandoned. The participants were not asked to rate how confident they were of the answers they gave. However, multiple words and phrases in the TA such as, “maybe”, “like”, “kind of”, and “I suppose”, reveal that some participants were more or less insecure of their answers.

The idiom meaning determination task scores are presented below in table 6 in proportion to the maximum points:

Table 6. Idiom meaning determination task scores

Participant	English	Swedish
A	2 / 14	3 / 12
B	10 / 14	4 / 8

C	10 / 14	7 / 14
D	7 / 14	2 / 14
E	7 / 14	2 / 10
F	11 / 14	4 / 8
G	12 / 14	1 / 10
H	7 / 14	4 / 8
I	8 / 14	6 / 8
all	74 / 126	33 / 92

As can be seen from the table, the maximum scores for Swedish and English differ, since the participants gave fewer answers for Swedish idioms. In other words, they managed to discover the meaning of less idioms in Swedish in the time given, an issue which will be more discussed in section 10.2. The overall score percentages were 58.7% for English, and 35.8% for Swedish. Thus, there were significant differences in the scores for English idioms when compared to Swedish idioms, with scores for English being slightly over 20 percentage points higher than the scores for Swedish. For some participants, such as participant G, this difference was extremely big. In contrast, when examining the scores of participants C and F, the differences level out a bit. Surprisingly, participants A and I succeed better with Swedish idioms. Whereas participant A's performance is equally low in both languages due to mainly employing guessing strategies, participant H proves that with the strategy of "using English", she was able to reach the highest score in Swedish. Participant H also had high chances of success in English as well. Nevertheless, leaving out parts of monolingual dictionary definitions turned out to be detrimental for full success in the task. Additionally, she seems to have made a mistake by defining the idiom *cut no ice* as its opposite: "make an impact", leaving out the word for negation.

What can account for these results? One clear pattern that can be observed from the data is that relying mostly on guessing often did not result in the correct meaning of the idiom. In contrast, when the participants consulted a dictionary, the answer was far more likely to be partially or fully correct. These findings are in agreement with Szczepaniak's (2006) findings on the inefficiency of guessing idioms from context compared to dictionary use. However, this does not mean that dictionary use guaranteed success, since on multiple occasions only parts

of the dictionary were used, information was dismissed, or erroneously reinterpeted. Moreover, some dictionaries provided erroneous meanings. Thus, success in the idiom determination task required skillful use of a dictionary, ability to extract the correct meaning, and evaluate the reliability of information.

Certain idioms seemed to be easier than others since they were known more often. These include the English idioms *hot air, play games, jump the gun* and *cut no ice*, and Swedish idiom *äga rum*. Perhaps one of the most reasonable explanations for this that is also evident from the VSC is that dictionary information was easily accessible of these otherwise quite opaque idioms. For all the English idioms, the translation offered by Google was almost always the first search result. Accessibility of information is also related to frequency, which likely explains the correct answers for the idiom *äga rum*. Of all the idioms chosen, it was the one with the most hits on Google. It could additionally be suspected that the participants had encountered the idioms before even if they had been identified as unfamiliar.

9.2 Error categories

The participant made several types of errors. The following idiom variation of form were found from the data; *bent over backwards, making noises, had it in for*, and *jump a gun.* These are evidence of the difficulties associated with idiom variations (see also Szczepaniak 2006), nonetheless these errors were not sanctioned since this study focused on meaning. Most of the errors related to meaning. In what follows, the error categories will be described and exemplified. The partially and fully incorrect answers were categorized into four categories. These were (1) literal meaning, (2) juxtaposition, (3) literal + non-literal meaning, and (4) non-literal meaning, which is further divided into a) non-literal meaning from context, b) non-literal meaning from other source, and c) Finnish idioms.

9.2.1 Literal meaning

Literal translations were partial or precise word-by-word Finnish translations of the idiom. Partial translations include answers such as "*itseään varten"* 'for/to oneself' for the idiom *göra rätt för sig*, of which only the last two words are translated. The majority of the answers were precise translations, such as "*toinen vitonen*" 'another five' for the Swedish idiom *en annan femma*. Most of the literal translations were produced by participants A and D, who interpreted most Swedish idioms literally. Ten out of eleven literal answers were Swedish idioms. In

conclusion, the participants seemed to be more prone to interpret Swedish idioms literally. The think-aloud data gives two explanations for this. Firstly, the D participant used *Google translate* for the idiom and directly copied the literal translation provided by it. This characterized five of participant D's answers. Her over-reliance on the translator proved to be detrimental for success in the task. Secondly, the participants may have translated the idiom to Finnish themselves, either directly or word by word with the help of a dictionary. For instance, participant A looked up multiple single words in a bilingual dictionary to form the translation of the idiom. Surprisingly, even if many of the participants tested the appropriateness of the literal translations in context, resulting in rather peculiar meanings, they saw no problem with the literal translation. Instead, the meaning of the context seemed to be modified and re-interpreted to fit the literal meaning and not vice versa. Literal errors indicate an inadequate knowledge of idioms, and their non-literal meanings.

9.2.2 Juxtaposition

In three of the answers, the participants had clearly drawn a parallel between the idiom and part of the example sentence. In the examples found, the meaning of the idiom had been juxtaposed with the meaning of the words in the sentence. Participant G, for instance, defines *gå in i väggen* as "get sick", which very much resembles the word *sjukskriven* 'on a medical leave' that appeared in the example sentence. In another answer, participant G had translated a large chunk of the example sentence (*"voi ottaa rennosti, koska ei ole niin paljon mitä pitää tehdä"*) by participant G and given it as the definition for the idiom *mellan varven*. Participant E's answer *houkutella* 'entice, lure' to the idiom *hold out an olive branch* is slightly different since the link to the sentence is not that explicit. However, entice and lure are both synonymous to invite, a word that appeared in the sentence, and hence this answer was also counted as an example of juxtaposition. By using this strategy, these two participants showed that they on the one hand were able to draw parallels, a skill related to using context (cf. Ames 1966). On the other hand, the meaning of the words in context was equaled with the meaning of the idiom, which resulted in erroneous interpretations.

9.2.3 Literal + non-literal meaning

The eight answers in the literal + non-literal category, were a mixture of translation, and non-literal meaning. Some of the answers were partly correct, whereas others were fully incorrect. For example, the Swedish idiom *take the edge off something* was defined as *"ottaa pahimman terän pois*" ('take the worst edge off'). This odd combination was taken from one translated

example sentence from the online context dictionary *Linguee*. An example of a fully incorrect answer of the idiom is *se om sitt hus* is "*riittää, että pitää huolta omasta kodistaan*" ('it is enough to take care of one's own home'). The first word in the definition is non-literal whereas the rest is translation. Participant C was most prone to give these types of answers. In three of her answers for Swedish idioms, a clear pattern can be seen: one word is literal, and the rest is figurative which each time resulted in a partially correct answer. For instance, *bent over backwards* was interpreted as "*taipua jonkin vuoksi tekemään jotain* ('bend/stretch because of someone to do something'). Even if bend can be interpreted figuratively in Finnish, it has a somewhat different nuance to it as illustrated with its English translations: "yield", "give in", and "defer to". These examples show that the participants partially do realize that the idioms are not completely literal, yet they were not fully willing to abandon the translation. Although this strategy differs from completely literal idioms, it provides further evidence of the prevalence of literal meanings.

9.2.4 Non-literal meaning

The errors in the first subcategory of non-literal meanings, were answers that fit well in the sentence context, and which could most often be traced back to the guessing from context strategy. Eight out of nine answers in this category were English idioms. Particularly, participant A tended to overgeneralize the contextual meaning of English idioms. Therefore, the definitions of the idioms were erroneous despite them being possible meanings if solely inspected in the context they occurred in. For instance, making noises was interpreted as reminding which made sense in the context *"His mother had started **making noises** about it being time for him to leave home"*. In addition, as many as three instances of the idiom take it on the chin could be found in this category. The sentence in which this idiom was embedded was the following: *"When the police arrived, **he took it on the chin** apologizing for the trouble he'd caused them*". As can be seen from the answers "realize your wrongdoing", "take responsibility", "admit your mistake", the sentence context has been utilized. The second answer was given by participant G who despite using several dictionaries overgeneralized the contextual meaning. The misleading effect of context in meaning determination has also been observed by Szczepaniak (2006) and Frantzen (2003).

The participants gave 12 Finnish idioms as fully or partially incorrect answers. Finnish idioms with partially or completely overlapping form and meaning were excluded from the study. In spite of this, many idioms similar in form or meaning were suggested. Most of the answers were formally similar idioms, i.e. false friends (see also Mäntylä 2004) that shared one

or two words with the target idiom. For instance, the Finnish idiom "pitää jalat maassa" ('keep feet on the ground') and *sätta ner foten* both have the word feet in common. Another answer for the same idiom was "ottaa härkää sarvista" ('take the bull by the horns'), which has vague resemblance to the actual meaning, "take action". One idiom that had semantic similarities with the target idiom *vara med på ett hörn*, was "*olla sivusta seuraajana*" ('be a bystander'). The correct and the partly erroneous meaning both convey a similar idea of not being fully/but only partially or less involved in something. But whereas the Finnish idiom implies passivity of the subject, the correct meaning entails being an actor that is doing something, albeit on a smaller scale. Even if each participant had their unique set of idioms, at times, the meaning of the same idiom was discovered multiple times. Interestingly, two exact same Finnish idioms were on multiple occasions suggested as the meaning of the idioms *in the loop* and *peka med hela handen.* The former was three times erroneously defined as *"kierteessä*" ('in a spiral'), and the latter was two times defined as "*näyttää kädestä pitäen*" ('show hands-on') although it in fact means "to instruct that something has to be done immediately". Many of these erroneous interpretations originated from bilingual/ized dictionaries, *Google translate*, and guessing. Still, the participants may have recognized that the target expressions could not be interpreted literally, and when they came up with a Finnish idiom, or saw it online, it was uncritically accepted. The suitability of the meaning in context may also have supported the choice of Finnish idioms as answers.

The last category consists of miscellaneous non-literal answers which did not fit in any of the above categories. Most of these answers had been inspired by dictionary definitions or included parts of dictionary excerpts. Consequently, slightly over half of the 22 answers were partially correct. Participant I was particularly inclined to leave parts of a monolingual dictionary definition out. As much as six of her answers for English idioms were only scored with one point. Below in figure 13 is a of screenshot of *Cambridge Dictionary's* definition for the idiom *on the ropes*:

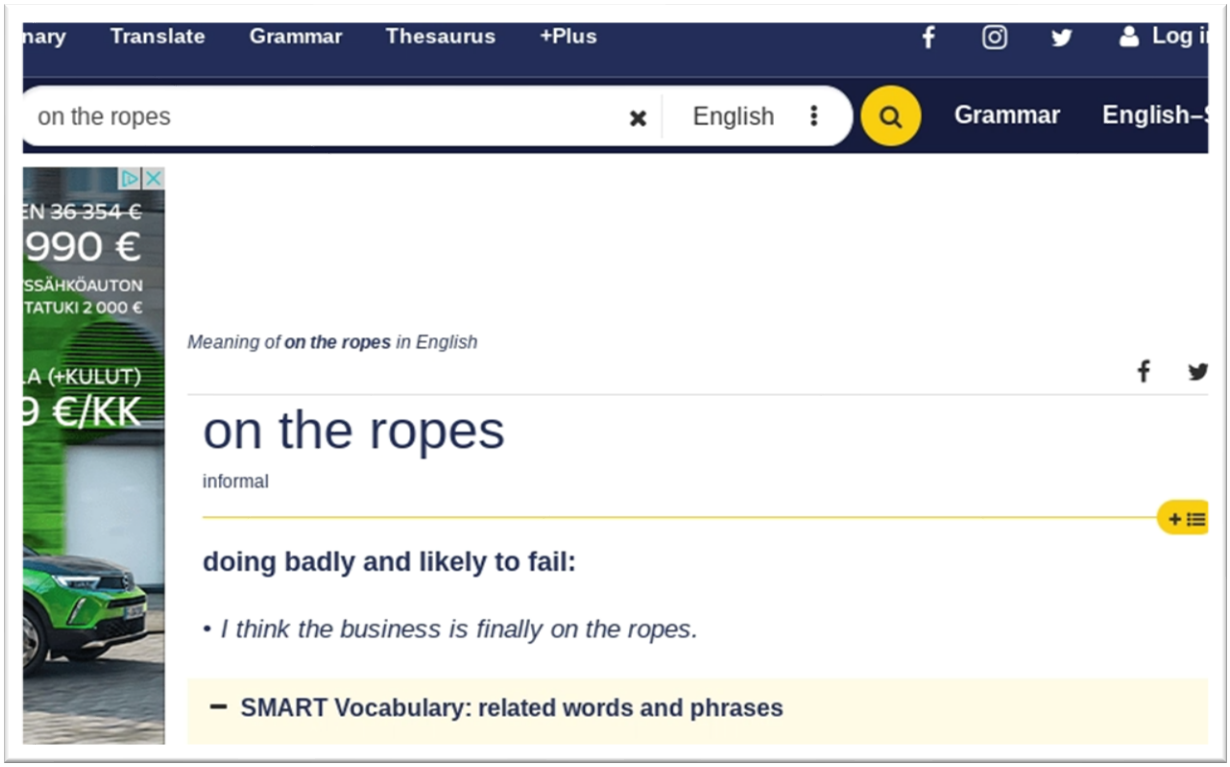

Figure 13. Cambridge Dictionary definition for the idiom "on the ropes"

After the dictionary consultation, participant I ended up giving "*jollakin menee huonosti*" ('someone is doing badly') as the definition. Thus, she had only taken the first part of the definition leaving out the last part. A similar pattern can also be observed regarding several other answers. Furthermore, there were erroneous answers that could be traced back to erroneous translations given by dictionaries. Certain answers seemed to be the result of illogical guesses that made no sense in the context. For example, *en annan femma*, which means another thing was defined as "*voisi tarkoittaa, että raha tulee ilmaiseksi hänelle tavallaan*" ('could mean that the money kind of comes freely to him'). This definition in fact contradicts the context, since buying is mentioned in the example sentence.

9.2.5 Summary of error categories

Considering the fact that the participants could use any available means for discovering the meaning of these unfamiliar idioms, the percentages are low, particularly for Swedish idioms. Figure 14 depicts the number of all error types:

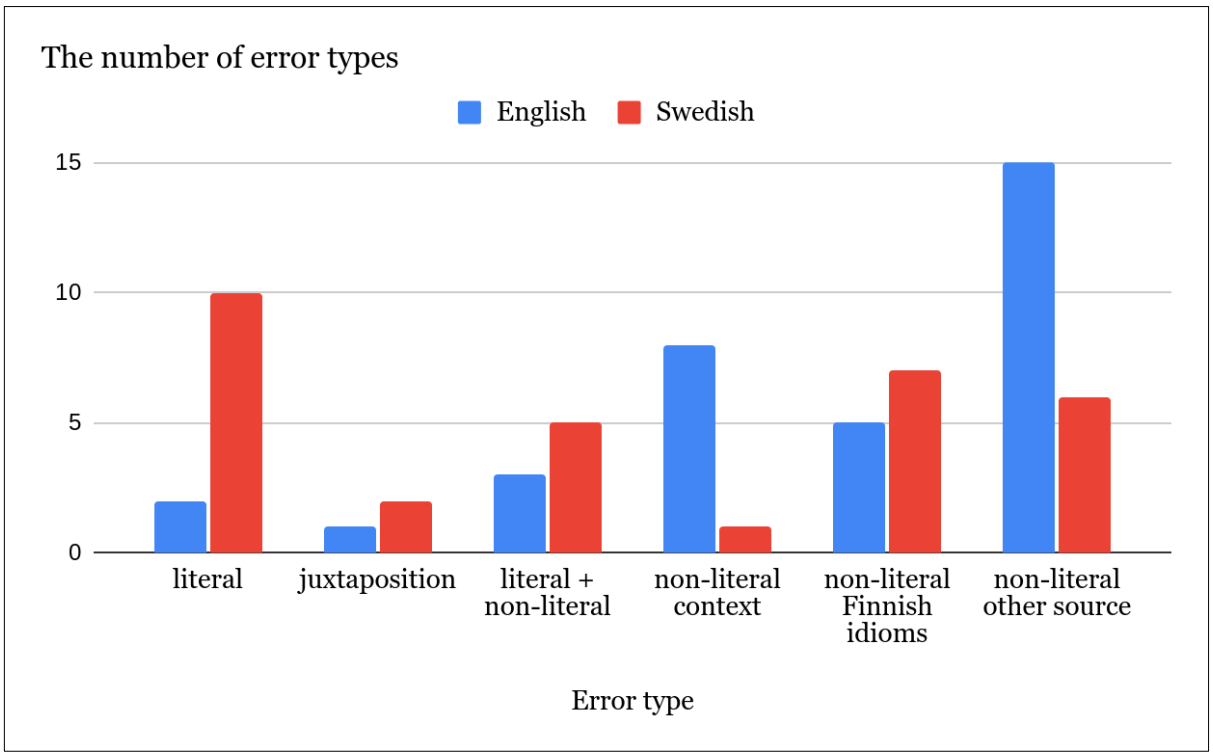

Figure 14. The number of error types

To summarize, error types (4c) non-literal meaning from other sources, (1) literal meaning, and (4b) Finnish idioms, were most common. If partly or completely literal (1, 3) answers are contrasted with completely figurative (4) answers, a conclusion can be drawn that non-literal meanings were preferred more. Differences between English and Swedish were also found. Swedish idioms were interpreted more literally, whereas answers for English idioms contained more overgeneralizations of context. However, the higher number of English idioms in the study may have had an influence on the results, and therefore, the findings should be interpreted with caution.

10. OTHER COMPARISONS BETWEEN ENGLISH AND SWEDISH

In this section, additional data of English and Swedish idioms will be compared to answer research question three.

10.1 Task duration

The maximum duration the participants were allowed to use for the task was 20 minutes per language. However, comparisons revealed that some participants completed the task in less time. The task durations are presented below in figure 15:

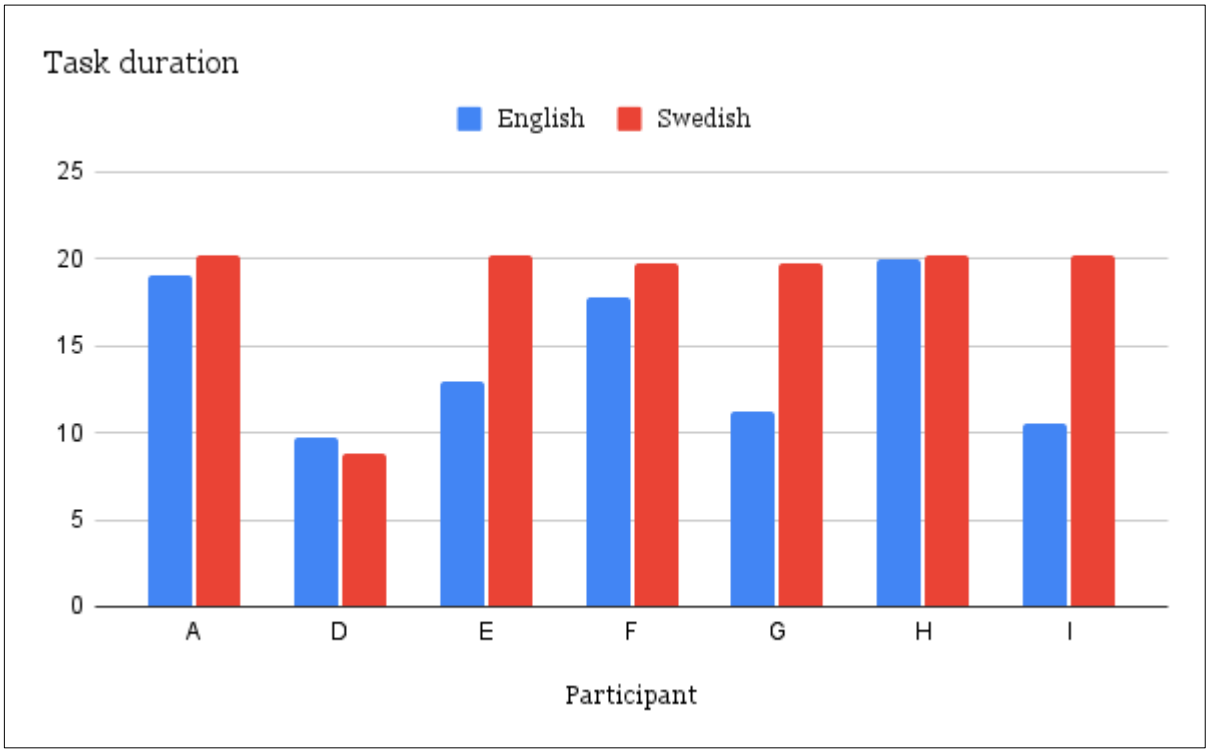

Figure 15. Task duration

It can be seen from figure 15 that more time was used for Swedish idioms than for English idioms. The mean for English was roughly 14 minutes and 30 seconds, in contrast to the mean for Swedish which was roughly 18 minutes and 30 seconds. However, if the scores of participant D, who seems to diverge from other participants in several different ways, are removed from the count, the mean for Swedish increases to 20 minutes and 5 seconds, which is close to the maximum time. Participant D performed tasks in both languages faster than any

other participant. The TA and the VSC together show that she did not dwell long on any of the idioms, which could hypothetically be attributed to efficiency, false confidence, or low motivation. All in all, discovering the meaning of Swedish idioms took 4-5,5 minutes longer than discovering the meaning of English idioms. Several explanations, including language proficiency, availability of information, strategies used, and the type of idiom may account for these results. Additionally, these findings give an explanation for the significantly larger number of strategies employed for Swedish, since more strategic actions could be employed in a larger amount of time.

10.2 Number of idioms

Figure 16 below presents the number of idioms discovered, and compares the number of Swedish and English idioms:

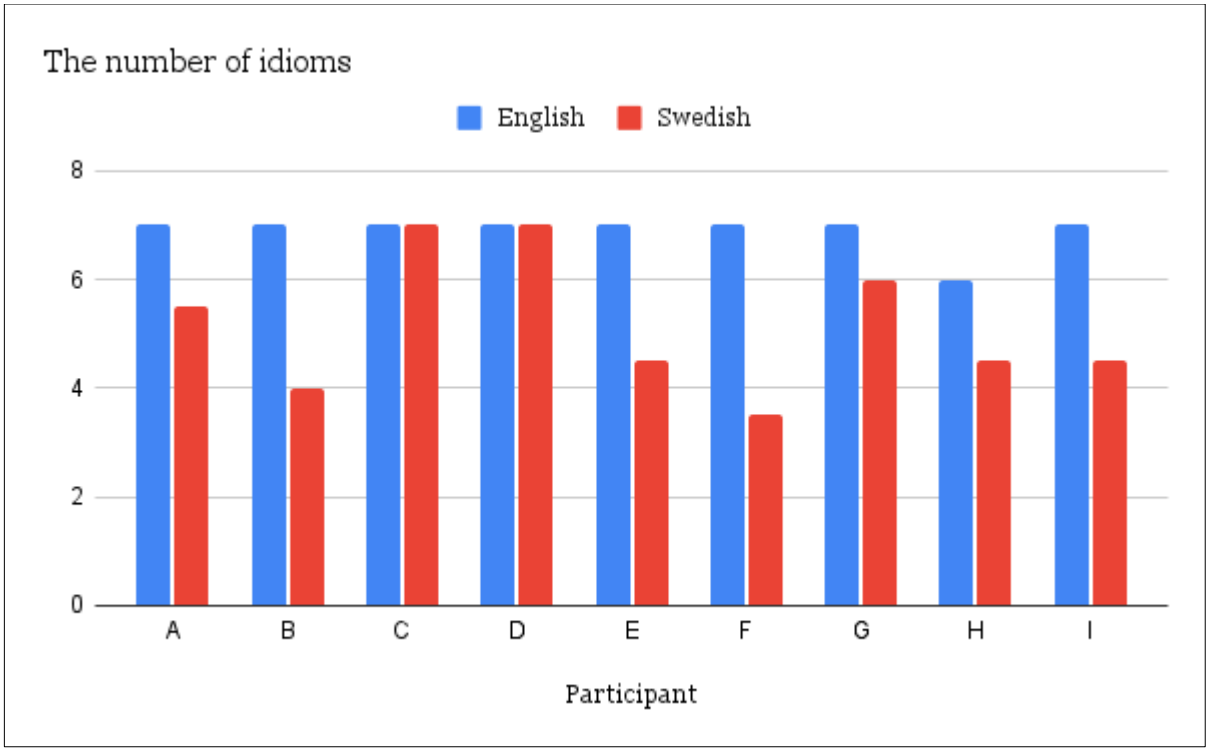

Figure 16. The number of idioms

All nine participants, except for participant H, managed to discover the meaning of the seven target English idioms. On the contrary, the mean number of the Swedish idioms was roughly five, with only two participants (C and D) out of nine being able to discover the meaning of all

Swedish idioms. Participant F, had the lowest number of Swedish idioms, discovering the meaning of as little as 3,5 idioms. Furthermore, as pointed out earlier, participant D seemed to be an exception in many other senses as well. The results indicate that the maximum time, approximately 3 minutes per each idiom, was not enough for discovering the meaning of Swedish idioms. All in all, results show that the participants were able to discover more English idioms in less time when compared to Swedish. As for idioms, the results suggest that discovering the meaning of idioms is laborious and time-consuming, particularly for Swedish.

11. DISCUSSION

In this section the findings will be discussed in light of previous research. To begin with, I will address the answers to the research questions one by one.

R1: What vocabulary learning strategies do Finnish upper secondary students use in discovering the meaning of English and Swedish idioms while thinking aloud?

Altogether 11 strategy categories used in the meaning discovery of idioms were found from the think-aloud protocols and the video screen capture recordings. These were: (1) selecting idioms, (2), assessing knowledge of a word or an mwu, (3) repetition of idioms, (4) translation, (5) guessing, (6) using a search engine (googling), (7) monolingual dictionary, (8) bilingual dictionary, (10) using other websites, (11) using English, and (12) other strategies. Most of the strategies have also been found in previous VLS studies on single words (e.g. Fan 2003; Gu & Johnson 1996; Pavicic-Takac 2008; Schmitt 1997), and idioms (Cooper 1999). However, the use of some strategies such as "analyzing part of speech", and "analyzing affixes and roots" from Schmitt's (1997) strategy taxonomy, was minimal probably since they are more helpful for learning single words. Using the literal meaning of the idiom, classified under the strategy of "guessing" was the only idiom specific strategy which was used enough to deserve mention. Other idiom specific strategies, such as pictorial elucidation and/or etymological elaboration (see e.g. Boers 2004; 2007; Karlsson 2019; Szczepaniak & Lew 2011; Ramonda 2016), and using metaphor (e.g. Beréndi et.al. 2008) were conspicuous with their absence. This may be explained by the nature of these strategies; they may be less used independently (Skoufaki 2005) and require preparations or instructions from a teacher. All in all, the results indicate that many VLS for single words can also be used to discover the meaning of idioms.

Calculations of the strategy instances revealed that bilingual dictionary use, guessing strategies, using a search engine (googling), and monolingual dictionary use were among the four most used strategies. These results are mainly consistent with previous studies. Alongside these more traditional strategies, some new strategies that have received less prominence in research literature, emerged. These include googling and using other informal web sources. Moreover, strategies were found to be very learner specific as learners differed both in the quality and quantity of the strategies used.

R2: How well are students able to discover the meaning of English and Swedish idioms? What types of errors do they make?

The overall task scores were 58.7% for English idioms and 35.8% for Swedish idioms. The percentages for both particularly English were higher when compared with studies that have investigated idiom knowledge (Macis & Schmitt 2017; McGavigan (2009 as cited in Milton 2009: 151-155). This is logical, since the participants could use any possible resources to discover the meaning of the idioms, instead of only relying on previous knowledge. The scores for English idioms were also found to be significantly higher than scores obtained from guessing the meaning of idioms from context (Szczepaniak's 2006: 75; Katsarou 2010: 288-289), and the scores from guessing + dictionary use (Szczepaniak 2006: 75). The former difference can be explained with dictionary use, and latter difference with the use of canonical forms of idioms instead of idiom variations, and access to a wider range of dictionary and web sources. Surprisingly, the 35.8% score for Swedish idioms, resembles the scores from Katsarou's (2010: 288-289) study (37.29%), and test results of Macis and Schmitt's (2017) study (33.02%). It is as if the benefits of being allowed to use a wider variety of VLS is nonexistent. From this point of view, the success rate for Swedish idioms is not flattering. Rather it corroborates earlier studies which have time and time again proved that learners struggle with learning idioms.

Four error types were discovered from the data. These were (1) literal meaning, (2) juxtaposition, (3) literal + non-literal meaning, and (4) non-literal meaning, which was further divided into a) non-literal meaning from context, b) non-literal meaning from other source, and c) Finnish idioms. Similar results have been obtained from previous studies. The learners in Mäntylä's (2004) study were also prone to accept literal interpretations of idioms, and Finnish false friends. Moreover, negative dictionary strategies and the misleading effect of context have also been observed in several studies investigating both idioms and single words (Frantzen 2003; Macis 2018; Szczepaniak 2006; Wingate 2004). In sum, some of the error categories were idiom-specific, whereas other error categories were more related to the type of strategy used.

R3: Are there any differences between English and Swedish?

Several differences between English and Swedish were observed in the data. Firstly, Finnish upper secondary students differed both quantitatively and qualitatively in their strategy use

depending on the language. Participants used a bilingual dictionary much more extensively for the meaning discovery of Swedish idioms, whereas a monolingual dictionary was more consulted for English idioms. Interestingly, one strategy specific for Swedish idioms was discovered. It involved using English language in order to search information about Swedish idioms. These findings provide supporting evidence for the language differences in strategy use obtained by Pavicic-Takac (2008: 134-145). However, most strategies seemed to be used for both languages, and could be hence regarded as "core VLS" (Pavicic-Takac 2008: 144). The comparisons of the meaning determination task scores also revealed a significant language difference; the overall scores for English idioms (58.7%) were much higher than those of Swedish (35.8%). Language also seemed to play a part in the type of errors learners committed. With Swedish idioms, learners were more prone to arrive at partially or fully literal answers. In contrast, with English idioms, overgeneralizations of non-literal contextual meaning were more common besides non-literal answers with their origin in other sources, such as dictionary use. In addition, the participants were able to discover the meaning of a higher number of English idioms in less time when compared to Swedish. Thus, it can be concluded that Swedish idioms were more difficult for learners than English idioms.

Many reasons may explain these differences. Firstly, there are far more speakers of English than speakers of Swedish. Additionally, English has a status of an international Lingua Franca. For these reasons, there may be more available online information, including dictionary information, about English idioms than Swedish ones. Secondly, Swedish proficiency is usually lower than English proficiency, which may have several consequences. For example, due to a lack of comprehension, the context may not be used to full extent. This may also be seen from dictionary definitions that were often arbitrary in the context. Dictionary use may also have been affected by proficiency since the comprehension of monolingual dictionary entries requires higher language skills.

This study has some limitations which will be addressed here. Firstly, this study only looked at meaning determination strategies, and hence, does not provide information about memorization/consolidation strategies. Secondly, the study did not measure how often and how well participants tend to use these strategies but what strategies the participants use when they face a high-stakes idiom meaning discovery task in an unusual situation. For these reasons, the results cannot directly be transferred to real life. However, if the participants demonstrated the use of a strategy, it shows that in real life the learner can use this VLS. Thirdly, the results are representative of a small homogenous group of Finnish Upper Secondary students. Nevertheless, the large number of idioms in this study allow making some generalizations.

Fourthly, the relationship between the VLS and its effect on the success in the idiom determination task was not investigated. Lastly, it was not studied how well the participants would have recognized or recalled the meaning and/or form of the idioms whose meanings were discovered correctly in this study.

There are also some factors that may have affected the results. Firstly, probably one of the most prominent factors was the fluctuation in how much time was used per language. It may have distorted the results regarding the idiom scores, error types, and strategy use. But despite this shortcoming, certain consistent patterns observed for both English and Swedish support the validity of these results. Secondly, the setting, for example having a laptop on display and the lack of interlocutors, may have primed the use of some VLS, and disfavored the use of other VLS, such as social strategies (Schmitt 1997). Likewise, awareness of the screen capture video may have primed the use of computer-related strategies. Thirdly, both English and Swedish idioms chosen for this study were among the more "high-frequent" idioms. However, there might have been variation among the frequency of English idioms, and there certainly was striking variation among the frequency of Swedish idioms (see Appendix 7). For instance, the Swedish idiom *äga rum* had over a million hits on Google, whereas other idioms rarely exceeded 150 000 hits. Even if frequency did not seem to have an effect in studies where idiom knowledge was tested (Karlsson 2012, Macis & Schmitt 2017; Mäntylä 2004), it may be a central criterion for including an idiom in a dictionary, and therefore affect the availability of information. This could be seen with the idiom *äga rum*, as it was three times defined fully correctly. Nevertheless, other similar observations from the data were not made. Furthermore, the same idioms were on multiple occasions defined correctly and/or partially incorrectly, which suggests that frequency after all, had a minor effect. Fourthly, particularly the Swedish example sentences proved to be difficult at times since some participants translated almost every word to Finnish. Even if guessing from context may not have been feasible due to a scarcity of contextual clues (cf. Webb 2008b), or a high density of unknown words (cf. Liu & Nation 1985), a more plausible explanation would be lower proficiency and vocabulary knowledge in Swedish, which have also been found to affect guessing from context (e.g. Nassaji 2006; Zahar, Cobb, & Spada, 2001). These methodological considerations should be taken into account when interpreting the results.

This study also produced interesting results regarding the video screen capture and the think-aloud method. Despite some of the problems related to these methods, all in all, this study has shown that extremely rich and versatile data can be gathered with these methods. However, some issues encountered will be addressed here. First and foremost, one possible downside of

a think-aloud task is learners' abnormal behavior (Nation 2013: 334), i.e. the observer's paradox. However, in this study my role as a researcher was only to occasionally prompt the learners to think aloud, and therefore it can be argued that interference on thinking or the behavior of the participants was minimal. Participants differ in their ability to vocalize thoughts (Ericsson and Simon 1993: 250), something which was observed in this study. Additionally, students may not be used to thinking aloud (Gu et.al. 2005: 289). For instance, one participant (D), explicitly commented after the training session that she had never done thinking aloud before. Nevertheless, unfamiliarity with the TA was prevented with instructions, training, and prompts However, the warm-up task of describing a picture, may have been too easy and too different from the real task, and hence minimized the transfer of TA into the performance of the real task.

The participants sometimes forgot to think aloud and some pauses occur in the data despite regular prompts. This may have been the result of the complexity of the task that in most cases involved talking, reading, typing on the laptop and writing. It is possible that being too engrossed in these activities lead to forgetting to verbalize (Gu et.al. 2005: 283). When there were longer pauses in verbalizing, prompts, such as "What are you thinking now?" were used. However, not all prompts had the desired effect. Instead of verbalizing, some participants answered the question by making a general comment, such as: "I'm thinking about the sentence", which was not helpful in eliciting strategy use. In these cases, the prompts may have interfered with thought processes since a specific answer was formulated for the researcher. Consequently, different prompts were used, such as "Keep thinking aloud". A similar problem was present throughout the TA of one participant, who seemed to be directing the speech at a listener despite instructions to verbalize "stream of thoughts" without having to address the speech to the researcher. An informal conversation with one participant also revealed that the TA method was insufficient in capturing the use of all strategies that were not verbalized (cf. Gu et.al. 2005: 289). On the contrary, dictionary use was always captured with VSC, which may have biased the amount of strategy use. Furthermore, if the information that was read from the screen was not read aloud and if it was not pointed with a cursor, strategy use could not be proved. Thus, to ensure that all strategy use is recorded, retrospective reports or interviews/ questionnaires could be used as a complement to TA protocols and VSC recordings. Additionally, the difficulty of the task should be carefully controlled, hence that the task does not require excessive concentration or effort (Gu et.al. 2005: 286). These issues should be considered in future studies.

Two broad pedagogical implications arise from the findings of this study. Firstly, as recommended by some scholars (Nation 2013; Gu 2019; Wingate 2004: 11) and encouraged by the benefits obtained from several studies (Craigo et.al. 2017; Mizumoto & Takeuchi 2009; Ranalli 2013; Zhao 2009) L2 learners should, regardless of the target language, be offered instruction and/or training in strategy use. Numerous observations from the data give cause to this. On multiple occasions, learners showed lack in skills regarding guessing from context, and dictionary use. For example, regarding dictionary use, participants often settled for first results of a google search or a dictionary look-up, misinterpreted dictionary definitions, read dictionary entries carelessly, and used the negative kidrule strategy (see also Chen 2016; Wingate 2004). Moreover, the growing trend of using various online information for meaning discovery, necessitates training in evaluation skills and critical media literacy since some of the information found online is of low quality (cf. Moon 2015; Nesi 2012). Secondly, learners should be familiarized with the basics of formulaic language and figurative language. This would also help them with the core features of idioms, including their recognition as multi-word units that are lexicogrammatically invariable, non-compositional, non-literal, often metaphorical in meaning. In addition, awareness should be raised about L1-L2 differences. Even if some participants demonstrated understanding of idioms, proneness to approaching idioms literally, using idiom variations in google and dictionary searches, and giving erroneous L1 idioms as interpretations, denotes a need for training. All in all, skills in the use of different strategies would not only benefit the learning of mwus but vocabulary learning in general. Moreover, familiarity with formulaic language and strategy skills may compensate for gaps in language proficiency.

Further similar research on VLS and idioms are required to establish the viability of the findings obtained from this study, particularly regarding language differences. In future investigations, one interesting research question that could be asked is what strategies learners actually use for learning mwus and idioms since studies thus far have not yet looked at learning mwus in a natural setting. Another question that could be investigated is the immediate and delayed learning effects of using specific strategies or strategy clusters for discovering the meaning of mwus or idioms, as this question has mostly been studied for memorization strategies (e.g. Beréndi et.al. 2008; Boers et.al. 2007; Ramonda 2016; Szczepaniak & Lew 2011). Most memorization strategies have proven to be useful tools for learning idioms, however, it is not known how applicable these strategies are in the Finnish context, and whether they apply to not just English but also other languages, such as Swedish.

12. REFERENCES

Abrahamsson, N. & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59(2), 249–306.

Alasuutari, P. (2007). *Laadullinen tutkimus* (3rd edition). Tampere: Vastapaino.

Ames, W. S. (1966). The development of a classification scheme of contextual aids. *Reading Research Quarterly*, 2(1), 57-82.

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (ed.), *Comprehension and teaching: Research reviews*. Newark, DE: International Reading Association, 77-117.

Atkins, B. T. S. & Varantola, K. (1998a). Language learners using dictionaries: The final report on the EURALEX/AILA research project on dictionary use. In, S. B. T. Atkins (ed.), *Using Dictionaries*. *Studies of Dictionary Use by Language Learners and Translators*. Tübingen: Max Niemeyer Verlag, 22–81.

Atkins, S.B. T. & Varantola, K. (1998b). Monitoring dictionary use. In Atkins, S. B. T Atkins (ed.), *Using Dictionaries. Studies of Dictionary Use by Language Learners and Translators*. Tübingen: Niemeyer Verlag, 83-122.

Arden-Close, C. (1993). NNS readers' strategies for inferring the meanings of unknown words. *Reading in a Foreign Language*, 9(2), 867-893.

Aydin, B. (2017). The effect of mental imaging technique on idiom comprehension in EFL learners. *International Journal of Languages' Education and Teaching*, 5(1), 201-212.

Bagheri, M. S. & Fazel, I. (2010). Effects of etymological elaboration on the EFL learners' comprehension and retention of idioms. *Journal of Pan-Pacific Association of Applied Linguistics,* 14(1), 45-55.

Barcroft, J. (2009). Strategies and performance in intentional L2 vocabulary learning. *Language Awareness*, 18(1), 74-89.

Barcroft, J. (2015). *Lexical Input Processing and Vocabulary Learning*. Amsterdam: John Benjamins Publishing Company.

Barkaoui, K. (2011). Think-aloud protocols in research on essay rating: An empirical study of their veridicality and reactivity. *Language Testing*, 28(1), 51-75.

Bengeleil, N. F. & Paribakht, T. S. (2004). L2 reading proficiency and lexical inferencing by university EFL learners. *Canadian Modern Language Review*, 61(2), 225-250.

Beréndi, M., Czábi, S. & Kövecses, Z. (2008). Using conceptual metaphors and metonymies in vocabulary teaching. In F. Boers & S. Lindstromberg (eds.), *Cognitive*

Linguistic Approaches to Teaching Vocabulary and Phraseology. Berlin; New York: Mouton de Gruyter, 65-99

Bishop, H. (2004). The effect of typographic salience on the look up and comprehension of unknown formulaic sequences. In N. Schmitt (ed.), *Formulaic sequences*. Amsterdam; Philadelphia: John Benjamins Publishing Company, 227-248

Boers, F. (2019). Factors affecting the learning of multiword items. In S. Webb (ed.), *The Routledge Handbook of Vocabulary Studies.* London; New York: Routledge, 143-157.

Boers, F., Demecheleer, M. & Eyckmans, J. (2004). Etymological elaboration as a strategy for learning idioms. In Bogaards, P. & Laufer, B. (eds.), *Vocabulary in a Second Language: Selection, Acquisition and Testing*. Amsterdam; Philadelphia: John Benjamins Publishing Company, 53-78.

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, H. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research*, 10, 245–261.

Boers, F., Eyckmans, J. & Stengers, H. (2007). Presenting figurative idioms with a touch of etymology: more than mere mnemonics? *Language Teaching Research*, 11(1), 43-62.

Boers, F. & Lindstromberg, S. (2008). How cognitive linguistics can foster effective vocabulary teaching? In F. Boers & S. Lindstromberg (eds.), *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology*. Berlin; New York: Mouton de Gruyter, 1-61.

Boers, F., Lindstromberg, S., Littlemore, J., Stengers, H. & Eyckmans, J. (2008). Variables in the mnemonic effectiveness of pictorial elucidation. In F. Boers & Lindstromberg, S. (eds), *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology*. Berlin: Mouton De Gruyter, 189-216.

Boers, F. & Lindstromberg, S. (2009). *Optimizing a lexical approach to instructed second language acquisition*. Hampshire; New York: Palgrave Macmillan.

Boers, F., Piquer Píriz A. M., Stengers, H. & Eyckmans, J. (2009). Does pictorial elucidation foster recollection of images? *Language Teaching Research*, 13(4), 367-382.

Bolander, M. (1989). Prefabs, patterns and rules in interaction? Formulaic speech in adult learners' L2 Swedish. In K. Hyltenstam & L. K. Obler (eds.), *Bilingualism across the lifespan: Aspects of acquisition, maturity, and loss*. Cambridge: Cambridge University Press, 73-86.

Brown, R., Waring, R. & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, 20(2), 136-163.

Chan, A. W. Y. (2012). A comparison between COBUILD, LDOCE5 and CALD3: Efficacy and effectiveness of the dictionaries for language comprehension and production. In R. Fjeld & J. Torjusen (eds). *Proceedings of the 15th EURALEX International Congress. 7-11 August 2012*. Oslo: Department of Linguistics and Scandinavian Studies, 606-612.

Chen, C. & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics*, 31(5), 693-713.

Chen, Y. (2016). Dictionary use for collocation production and retention: A CALL-based study. *International Journal of Lexicography*, 30(2), 225-251.

Cieślicka, A. (2006). Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research*, 22(2), 115-144

Clarke, D. F. & Nation, I. S. P. (1980). Guessing the meanings of words from context: strategy and techniques. *System*, 8(3), 211-220.

Cohen, A. D. (2013). *Verbal report.* In C. A. Chapelle (ed.), *The Encyclopedia of Applied Linguistics*. Oxford: Wiley-Blackwell.

Cohen, A. D. & Wang, I. K-H. (2018). Fluctuation in the functions of language learning strategies. *System* 74, 169-182.

Collins Cobuild Idioms dictionary (4th edition). (2020). Hands, P., Moore, J., Maree, A., Todd, L. & Woods, S. (eds.). Glasgow: HarperCollins Publisher.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied linguistics*, 29(1), 72-89.

Cooper, T. C (1999). Processing of idioms by L2 learners of English. *TESOL Quarterly*, 33(2), 233-262.

Cowan, J. (2017). The potential of cognitive think-aloud protocols for educational action-research. *Active Learning in Higher Education*, 20(3), 219-232.

Craigo, L., Ehri, L. C. & Hart, M. (2017). Teaching community college students strategies for learning unknown words as they read expository texts. *Higher Learning Research Communications*, 7(1), 43-64.

Craik, F. I. M. & Lockhart, R. S. (1972). Levels of Processing: A framework for memory research. *Journal of verbal learning and verbal behavior*, 11, 671-684.

Daneman, M. & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, 25(1), 1-18.

Day, R. R., Omura, C. & Hiramatsu, M. (1991). Incidental EFL vocabulary Learning and Reading. *Reading in a Foreign Language*, 7(2), 541-551.

De Schryver, G., Joffe, D., Joffe, P. & Hillewaert (2006). Do dictionary dictionary users really look up frequent words? - On the overestimation of the value of corpus-based lexicography. *Lexikos*, 16, 67-83.

Durrant, P. & Schmitt, N. (2010). Adult learners' retention of collocation from exposure. *Second Language Research*, 26(2), 163-188.

Dziemianko, A. (2010). Paper or electronic. The role of dictionary form in language reception, production and the retention of meaning and collocations. *International Journal of Lexicography*, 23(3), 257-273.

Ellis, N. & Beaton, A. (1993). Factors affecting the learning of foreign language vocabulary: imagery keyword mediators and phonological short-term memory. *The Quarterly Journal of Experimental Psychology*, 46A(3), 533-558.

Ellis, N. C. (1995). The psychology of foreign language vocabulary acquisition: Implications for CALL. *International Journal of Computer Assisted Language Learning (CALL)*, 8(2-3), 103-128.

Enström, I. (2013). Ordförråd och ordinlärning - med särskilt fokus på avancerade inlärare. In K. Hyltenstam & I. Lindberg (eds.), *Svenska som andraspråk.* (2nd edition). Studentlitteratur, 169-195.

Ericsson, K. A. & Simon, H. A. (1993). *Protocol analysis: Verbal Reports as Data* (rev. ed.). Cambridge; London: MIT Press.

Erman, B. & Warren, B. C. (2000). The idiom principle and the open choice principle. *Text & Talk*, 20(1), 29-62.

Fan, M. (2000). How big is the gap and how to narrow it? An investigation into the active and passive vocabulary knowledge of L2 learners. *RELC Journal*, 31(2), 105-119.

Fan, M. (2003). Frequency of use, perceived usefulness, and actual use of second language vocabulary strategies: A study of Hong Kong learners. *The Modern Language Journal*, 87(2), 222-241.

Fischer, R. (2007). How do we know what students are actually doing? Monitoring students' behavior in CALL. *Computer Assisted Language Learning*, 20(5), 409-442.

Tutkimuseettinen neuvottelukunta (TENK) (2019). Hyvä tieteellinen käytäntö (HTK). Hyvä tieteellinen käytäntö (HTK) | Tutkimuseettinen neuvottelukunta (tenk.fi) (15 March, 2021).

Fox, M. C., Ericsson, K. A. & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin*, 137(2), 316-344.

Frantzen, D. (2003). Factors affecting how second language Spanish students derive meaning from context. *Modern Language Journal*, 87, 168-199.

Fraser, B. (1970). Idioms within a transformational grammar. *Foundations of Language*, 6(1), 22-42.

Gánem-Gutiérrez, G. A. & Gilmore, A. (2021). A mixed methods case study on the use and impact of web-based lexicographic tools on L2 writing. *Computer Assisted Language Learning.* 12 October 2021, 1-27.

Gu, P. Y. & Johnson, R. K. (1996). Vocabulary Learning strategies and language learning outcomes. *Language Learning*, 46(4), 643-679.

Gu, P. Y. (2002) Gender, academic major, and vocabulary learning strategies of Chinese EFL learners. *RELC Journal*, 33(1), 35-54.

Gu, P. Y. (2003a). Fine brush and freehand: The vocabulary-learning art of two successful Chinese EFL learners. *TESOL Quarterly*, 37(1), 73-104.

Gu, P. Y. (2003b). Vocabulary learning in a second language: Person, task, context and strategies. *TESLEJ*, 7(2). Retrieved from http://tesl-ej.org/ej26/a4.html

Gu, P. Y., Hu, G. & Zhang L. J. (2005). Investigating language learner strategies among lower primary school pupils in Singapore. *Language and Education*, 19(4), 281-303.

Gu, P. Y. (2010). Learning strategies for vocabulary development. *Reflections on English Language Teaching*, 9(2), 105-118.

Gu, P. Y. (2019). Strategies for Learning Vocabulary. In S. Webb (ed.), *The Routledge Handbook of Vocabulary Studies*. London; New York: Routledge, 271-287.

Grant, I. E., & Bauer, L. (2004). Criteria for redefining idioms: Are we barking up the wrong tree? *Applied Linguistics*, 25(1), 38–61.

Grant, L., E. (2006). Frequency of 'core idioms' in the British National Corpus (BNC). *International Journal of Corpus Linguistics*, 10(4), 429-451.

Grant, L. E. & Nation, P. (2006). How many idioms are there in English? *International Journal of Applied Linguistics*, 151(1), 1-14.

Haghshenas, M. S M. & Hashemian, M. (2016). A comparative study of the effectiveness of two strategies of etymological elaboration and pictorial elucidation on idiom learning. A case of young EFL Iranian learners. *English Language Teaching*, 9(8), 140-151.

Hamel, M. J. (2012). Testing aspects of the usability of an online learner dictionary prototype: a product- and process-oriented study. *Computer Assisted Language Learning*, 25(4), 339-365.

Hamel, M. J. & Séror, J. (2016). Video screen capture to document and scaffold the L2 writing process. In C. Caws & M. J. Hamel (eds.), *Language-learner computer interactions: Theory, methodology, and applications*. Amsterdam; Philadelphia: John Benjamins Publishing Company, 137-162.

Heid, U. & Zimmermann, J. (2012). Usability testing as a tool for e-dictionary design: collocations as a case in point. In R. Fjeld & J. Torjusen (eds), *Proceedings of the 15th EURALEX International Congress. 7-11 August 2012*. Oslo: Department of Linguistics and Scandinavian Studies, 661-671.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language acquisition*, 21(2), 303-317.

Hulstijn, J., Hollander, M. & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign-language students: the influence of marginal glosses, dictionary use, and reoccurence of unknown words. *The Modern Language Journal*, 80(3), 327-339.

Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty & M. H. Long (eds.), *The handbook of second language acquisition*. Blackwell Publishing, 349-381.

Hult, A. (2009). Kort och gott - Om idiomens språkprov i Svensk ordbok utgiven av Svenska Akademien. *Nordiska Studier i Lexikografi 10, 209-222*.

Hult, A., Malmgren, S. & Sköldberg (2010). Lexin - a report from a recycling lexicographic project in the North. In A. Dykstra & T. Schoonheim (eds.), *Proceedings of the XIV EURALEX International Congress. 6-10 July 2010*. Leeuwarden; Ljouwert: Fryske Akademy – Afûk, 800-809.

Hu, H. M. & Nassaji, H. (2014). Lexical inferencing strategies: The case of successful versus less successful inferencers. *System*, 45(1), 27-38.

Hu, M. & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403-430.

Hu, J. & Gao, X. A. (2017). Using think-aloud protocol in self-regulated reading research. *Educational Research Review*, 22, 181-193.

Hyltenstam, K (2016). *Advanced Proficiency and Exceptional Ability in Second Languages*. Boston: De Gruyter Mouton.

Hockett, C. F. (1958). *A Course in Modern Linguistics*. New York: MacMillan Company.

Irujo, S. (1986). Don't put your leg in your mouth: transfer in the acquisition of idioms in a second language. *Tesol Quarterly*, 20(2).

Jenkins, J. R. & Dixon, R. (1983). Vocabulary learning. *Contemporary Educational Psychology*, 8, 237-260.

Jiménez Catalán, R. M. (2003). Sex differences in L2 vocabulary learning strategies. *International Journal of Applied Linguistics*, 13(1), 54-77.

Karlsson, M. (2019). *Idiomatic Mastery in a First and Second Language.* Bristol: Multilingual Matters.

Katsarou, E. C. (2010). *The Use of Lexical Inferencing Strategies in the identification and comprehension of L2 Phrasal Idioms During Reading by Greek Learners of English.* University of Essex.

Kecskes, I. (2000). A cognitive-pragmatic approach to situation-bound utterances. *Journal of Pragmatics*, 32, 605-625.

Kim, C. (2016). L2 learners' recognition of unfamiliar idioms composed of familiar words. *Language Awareness*, 25(1-2), 89-109.

Kjellmer, G. (1996). Idiomen, kollokationerna och lexikonet. *LexicoNordica* 3, 79-91.

Knight, S. (1994). Dictionary use while reading: The effect on comprehension and vocabulary acquisition for students of different verbal abilities. *The Modern Language Journal*, 78(3), 285-299.

Kulikova, O. (2015). *Vocabulary learning strategies and beliefs about vocabulary learning: a study of beginning university students of Russian in the United States*. University of Iowa.

Kojic-Sabo, I. & Lightbown, P. M. (1999). Students' approaches to vocabulary learning and their relationship to success. *The Modern Language Journal*, 83(3), 176-192.

Kosem, I. et. al. (2018). The image of the monolingual dictionary across Europe. Results of the European survey of dictionary use and culture. *International Journal of Lexicography*, 32(2), 91-114.

Kovanen, M. (2014). *Vocabulary learning strategies employed by Finnish high school EFL students*. University of Jyväskylä.

Koyama, T. & Takeuchi, O. (2009). How effectively do good language learners use handheld electronic dictionaries: A qualitative approach. *Language education & Technology*, 46, 131-150.

Kövecses, Z. & Szabó, P. (1996). Idioms: A view from cognitive semantics. *Applied Linguistics*, 17(3), 326-355.

LaBontee, R. A. (2019). *Strategic Vocabulary Learning in the Swedish Second Language Context.* University of Gothenburg.

Laffey, D. (2020). Vocabulary learning strategies preferred by Korean University Students. *English Teaching*, 75(4), 81-100.

Lakoff, G. & Johnson, M. (1980). Conceptual metaphor in everyday language. *The Journal of Philosophy*, 77(8), 453-486.

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. (2011). The contribution of dictionary use to the production and retention of collocations in a second language. *International Journal of Lexicography*, 24(1), 29-49.

Laufer, B. & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics*, 29(4), 694-716.

Laufer, B. & Hadar, L. (1997). Assessing the effectiveness of monolingual, bilingual, and "bilingualised" dictionaries in the comprehension and production of new words. *The Modern Language Journal*, 81(2), 189-196.

Laufer, B. & Hill, M. (2000). What lexical information do L2 learners select in a CALL dictionary and how does it affect word retention. *Language, Learning and Technology*, 3(2), 58-76.

Laufer, B. & Melamed, L. (1994). Monolingual, bilingual and bilingualised dictionaries: Which are more effective, for what and for whom? In W. Martin, W. Meijs., E. P., P. Van Sterkenburg & P. Vossen (eds.), *Euralex 1994 proceedings 14*. Amsterdam: International Congress on Lexicography, 565-76.

Laufer, B. & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: effects of language learning context. *Language Learning*, 48(3), 365-391.

Lawson, M. J. & Hogben, D. (1996). The vocabulary-learning strategies of foreign-language students. *Language Learning*, 41(1), 101-135.

Lewis, M. (2012). *The lexical approach: The state of ELT and a way forward.* Milton Keynes: Cengage Learning EMEA.

Lew, R. (2004). *Which dictionary for whom? Receptive use of bilingual, monolingual and semi-bilingual dictionaries by Polish learners of English*. Doctoral dissertation. Poznan: Motivex.

Lew, R. (2012). How can we make electronic dictionaries more effective? In S. Granger & M. Paquot (eds.), *Electronic Lexicography*. Oxford: Oxford University Press, 343-362.

Lew, R. (2013). Online dictionary skills. *Proceedings of eLex 2013*, 16-31.

Lew, R. & Adamska-Salaciak, A. (2015). A case for bilingual learners' dictionaries. *ELT Journal*, 69(1), 47-57.

Lew, R. and De Schryver, G. (2014). Dictionary users in the digital revolution. *International Journal of Lexicography*, 27(4), 341-359.

Lew, R. & Doroszewska, J. (2009). Electronic dictionary entries with animated pictures: Lookup preferences and word retention. *International Journal of Lexicography*, 22(3), 239-257.

Li, J. (2010). Learning vocabulary via computer-assisted scaffolding for text processing. *Computer Assisted Language Learning*, 23(3), 253-275.

Li, L. & Xu, H. (2015). Using an online dictionary for identifying the meanings of verb phrases by Chinese EFL learners. *Lexikos 25*, 191-209.

Lindstromberg, S. (2019). Intentional vocabulary learning. In S. Webb (ed.), *The Routledge handbook of vocabulary knowledge*. London; New York: Routledge, 240-254.

Littlemore, J., Chen, P. T., Koester, A. & Barnden, J. (2011). *Applied Linguistics*, 32(4), 408-429.

Liu, D. (2008). *Idioms. Description, Comprehension, Acquisition, and Pedagogy*. New York: Routledge.

Liu, N. & Nation, I. S. P. (1985). Factors affecting guessing vocabulary from context. *RELC Journal*, 16(1), 33-42.

Liontas, J. (2003). Killing two birds with one stone: Understanding Spanish VP idioms in and out of context. *Hispania*, 86(2), 289-301.

Lukion Opetussuunnitelman perusteet (2019). Helsinki: Opetushallitus.

Luthman, H. (2017). *Svenska idiom: 5.000 vardagsuttryck* (3rd edition.). Lund: Folkuniversitetets förlag.

Lorentzen, H. & Theilgaard, L. (2012). Online dictionaries - how do users find them and what do they do once they have? In R. Fjeld & J. Torjusen (eds.), *Proceedings of the 15th EURALEX International Congress.* 7-11 August 2012. Oslo: Department of Linguistics and Scandinavian Studies, 654-660.

Luppescu, D. & Day, R. (1993). Reading, Dictionaries, and Vocabulary Learning. *Language Learning*, 43(2), 263-279.

Macis, M. (2018). Incidental learning of duplex collocations from reading: Three case studies. *Reading in a Foreign Language*, 30(1), 48-75.

Macis, M. & Schmitt, N. (2017). Not just 'small potatoes': Knowledge of the idiomatic meanings of collocations. *Language teaching research*, 21(3), 321-340.

Makkai, A. (1972). *Idiom Structure in English*. Hague; Paris: De Gruyter Mouton.

Martinez, R. & Murphy, V. A. (2011). Effect of Frequency and idiomaticity on second language reading comprehension. *TESOL Quarterly*, 45(2), 267-290.

Mayring, P. (2014). *Qualitative content analysis: Theoretical foundations, basic procedures and software solution.* Klagenfurt, Germany: SSOAR Open Access.

Melka, F. (1997). Receptive vs. productive aspects of vocabulary. In N. Schmitt & M. McCarthy (eds.), *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press, 84-102.

Mekheimer, M., A. (2018). Effects of e-dictionaries on reading comprehension and vocabulary learning in EFL college students: a re-examination. *Journal of the Faculty of Education at beni Suef*, 1, 413-464.

Mizumoto, A. (2010). *Exploring the art of vocabulary learning strategies: a closer look at Japanese EFL university students*. Doctoral dissertation. Tokyo: Kinseido.

Mizumoto, A. & Takeuchi, O. (2009). Examining the effectiveness of explicit instruction of vocabulary learning strategies with Japanese EFL university students. *Language Teaching Research*, 13(4), 425-449.

Mondria, J-A. & Wiersma, B. (2004). Receptive, productive, and receptive + productive L2 vocabulary learning: What difference does it make? In P. Bogaards & B. Laufer (eds.), *Vocabulary in a Second language*. Amsterdam; Philadelphia: John Benjamins Publishing Company, 79-100.

Moon, R. (1998). *Fixed Expressions and Idioms in English, A Corpus-Based Approach.* Oxford: Clarendon Press.

Moon, R. (2015). Idioms: a view from the web. *International Journal of Lexicography*, 28(3), 318-337.

Mäntylä, K. (2004). *Idioms and language users: The effect of the characteristics of idioms on their recognition and interpretation by native and non-native speakers of English*. University of Jyväskylä.

Nagy, W. E., Anderson, R. C. & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, 1 June 1987.

Nagy, W. E. & Scott J. A. (2000). Vocabulary Processes. In M. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr, (eds.), *Handbook of Reading Research*, Volume III. Muhwah, N.J.: Lawrence Erlbaum Associates, 269-284.

Nakata, T. (2019). Learning words with flashcards and word cards. In S. Webb (ed.), *The Routledge Handbook of Vocabulary studies*. London; New York: Routledge, 304-317.

Nassaji, H. (2003). L2 vocabulary learning from context: strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *TESOL Quarterly*, 37(4), 645-670.

Nassaji, H. (2006). The relationship between depth of vocabulary knowledge and L2 learner's lexical inferencing strategy use and success. *The Modern Language Journal*, 90, 387-401.

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd edition.). New York: Cambridge University Press.

Nation, P. (2019). The different aspects of vocabulary knowledge. In S. Webb (ed.), *The Routledge Handbook of Vocabulary Studies*. London; New York: Routledge, 15-29.

Nation, I. S. P. & Webb, S. (2011). *Researching and Analyzing Vocabulary*. Boston: Heinle, Cengage learning.

Nation, I. S. P. & Wang, K. M. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, 12(2), 355-380.

Nattinger, J. R. & DeCarrico, J. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.

Nesi, H. (1999). The specification of dictionary reference skills in higher education. In R.R.K. Hartmann (ed.), *Dictionaries in language learning. Recommendations, national reports and thematic reports from the Thematic Network Project in the Area of Languages, sub-project 9: dictionaries.* Berlin: Freie Universität Berlin, 53-67.

Nesi, H. (2012). Alternative e-dictionaries: Uncovering dark practices. In S. Granger & M. Paquot (eds.), *Electronic Lexicography*. Oxford: Oxford University Press, 363-378.

Niemi, J., Mulli, M., Nenonen, S., Nikolaev, A. & Penttilä, E. (2013). Idiomatic proclivity and literality of meaning in body-part nouns: Corpus studies of English, German, Swedish, Russian and Finnish. *Folia Linguistica*, 47(1), 237-242.

Noroozi, I. & Salehi, H. (2013). The effect of the etymological elaboration and rote memorization on learning idioms by Iranian EFL students. *Journal of Language Teaching and Research*, 4(4), 845-851

Oxford, R. L (1990). *Language learning strategies: what every teacher should know?* Boston: Heinle & Heinle.

Paivio, A. (1971). Imagery and language. *Current Cognitive Approaches* 1971, 7-32.

Parry, K. (1997). Vocabulary and comprehension: Two portraits. In J. Coady & T. Huckin (eds.), *Second language vocabulary acquisition*: *A rationale for pedagogy*, Cambridge: Cambridge University Press, 55-68

Pavia, N., Webb, S. & Faez F. (2019). Incidental vocabulary learning through listening to songs. *Studies in Second Language Acquisition*, 41(4), 745-768.

Pavicic-Takac, V. (2008). *Vocabulary learning strategies and foreign language acquisition*.Bristol, Blue Ridge Summit: Multilingual Matters.

Pawley, A. and F, Syder. 1983. Two puzzles for linguistic theory. In Richards, J. C. Richards & R. W. Schmidt (eds.) *Language and Communication*. New York; London: Longman, 191-226.

Pellicer-Sánchez, A. (2016). Incidental L2 vocabulary acquisition *from* and *while* reading. An eye-tracking study. *Studies in Second Language Acquisition*, 38(1), 1-34.

Pellicer-Sánchez, A. (2017). Learning L2 collocations incidentally from reading. *Language Teaching Research,* 21(3), 381-402.

Peters, E. (2007). Manipulating L2 learners' online dictionary use and its effect on L2 word retention. *Language Learning and Technology*, 11(2), 36-58.

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, 40(3), 551-577.

Pigada, M. & Schmitt, N. (2006). Vocabulary Acquisition from Extensive Reading: A case study. *Reading in a Foreign Language*, 18(1), 1-28.

Pulido, D. (2007). The effects of topic familiarity and passage sight vocabulary on L2 lexical inferencing and retention through reading. *Applied Linguistics* 28(1), 66-86.

Prentice, J. (2010). Käppen i hjulen. Behärskning av svenska konventionaliserade uttryck bland gymnasieelever med varierande språklig bakgrund. *Rapporter om svenska som andraspråk (ROSA)*. Institutet för svenska som andraspråk, Institutionen för svenska språket. Göteborgs universitet.

Prentice, J. & Sköldberg, E. (2010). Klättra på väggarna eller bara vara ett med soffan? Om figurativa ordförbindelser bland ungdomar i flerspråkiga skolmiljöer. *Språk & stil*, 20, 5-35.

Prichard, C. & Matsumoto, Y. (2011). The effect of lexical coverage and dictionary use on L2 reading comprehension. *Reading Matrix: An Inteernational Online Journal*, 11(3), 207-225.

Qian, D. D. & Lin, L. H. F. (2019). The relationship between vocabulary usefulness: high-, mid-, and low-frequency words. In S. Webb (ed.), *The Routledge Handbook of Vocabulary Studies*. London; New York: Routledge, 66-80.

Ramonda, K. (2016). *The role of semantic transparency and metaphorical elaboration through pictures for learning idioms in a second language.* University of Birmingham.

Ranalli, J. (2013). Online strategy instruction for integrating dictionary skills and language awareness. *Language, Teaching and Technology* 17(2), 75-99.

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language testing*, 10(3), 355-371.

Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly*, 10(1), 77-89.

Rodgers, M. P. H. & Webb, S. (2020). Incidental vocabulary learning through viewing television. *ITL - International Journal of Applied Linguistics*, 171(2), 191-220.

Sanaoui, R. (1995). Adult learners' approaches to learning vocabulary in second languages. *The Modern Language Journal*, 79(1), 15-28.

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (eds.), *Vocabulary; Description, acquisition and pedagogy*. Cambridge: Cambridge University Press, 199-227.

Schmitt, N. (2004). *Formulaic Sequences. Acquisition, processing and use*. Amsterdam; Philadelphia: John Benjamins Publishing Company.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual.* New York: Palgrave Macmillan.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language testing*, 18(1), 55-88.

Scholfield, P. (1982). Using English for Comprehension. *TESOL Quarterly*, 16(2), 185-194).

Schreier, M. (2014). Content analysis, qualitative. In U. Flick (ed.), *The SAGE Handbook of Qualitative Data Analysis*. London: Sage, 170-183.

Selvi, A. F. (2019). Qualitative content analysis. In J. McKinley & H. Rose (eds.), *The Routledge Handbook of Research Methods in Applied Linguistics*. London: Routledge.

Shefelbine, J. L. (1990). Student factors related to variability in learning word meanings from context. *Journal of Reading Behavior*, 22(1), 71-97.

Shin, D., Chon, Y. V. & Kim, H. (2011). Receptive and productive vocabulary sizes of high school learners: What next for the basic word list? *English Teaching*, 66(3), 127-152.

Skoufaki, S. (2005). Use of conceptual metaphors: A strategy for the guessing of an idiom's meaning? In M. Mattheoudakis & A. Psaltou-Joycey (eds.), *Selected Papers on*

Theoretical and Applied Linguistics from the 16th International Symposium, 11-13 April 2003, Aristotle University of Thessaloniki, 542-556.

Smith, P. & King, J. R. (2013). An examination of veridicality in verbal protocols of language learners. *Theory and Practice in Language Studies*, 3(5), 709-720.

Stengers, H., Boers, F., Housen, A. & Eyckmans, J. (2011). Formulaic sequences and oral proficiency: Does the type of target language influence the association? *International Review of Applied Linguistics in Language Teaching*, 49(4), 321-343.

Steinel, M., Huljstin, J. & Steinel, W. (2007). Second language idiom learning in a paired-associate paradigm: Effects of direction of testing, idiom imageability, and idiom transparency. *Studies in Second Language Acquisition*, 29(3), 449-484.

Szczepaniak, R. (2006). *The Role of Dictionary Use in the Comprehension of Idiom Variants.* Tübingen: De Gruyter. Lexicographica.

Szczepaniak, R. & Lew (2011). The role of imagery in dictionaries of idioms. *Applied Linguistics*, 32(3), 323-347.

Szudarski, P. & Carter, R. (2016). The role of input enhancement and input flood in the acquisitions of collocations by EFL learners. *International Journal of Applied Linguistics*, 26(2), 245-265.

Tarp, S. (2009). Reflections on lexicographical user research. *Lexikos*, 19, 275-296.

Tuomi, J. & Sarajärvi, A. (2018) *Laadullinen tutkimus ja sisällönanalyysi*. Helsinki: Tammi.

Van Zeeland, H. & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics*, 34(4), 457-479.

Vasiljevic, Z. (2015). Effects of etymology and pictorial support on the retention and recall of L2 idioms. *Electronic Journal of Foreign Language Teaching*, 12(1), 35-55.

Verlinde, S. & Binon, J. (2010). Monitoring dictionary use in the electronic age. In *Proceedings of the XIV Euralex International Congress*. Ljouwert: Afuk; 1144-1151

Waring, R. (1997). A Comparison of the Receptive and Productive Vocabulary Sizes of Some Second Language Learners, *Immaculata* 1, 53-68.

Waring, R. & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15(2), 130-163.

Webb, S. (2008a). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language acquisition*, 30(1), 79-95.

Webb, S. (2008b). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, 20(2), 232-245.

Webb, S. (2013). Depth of vocabulary knowledge. In C. A. Chapelle (ed.), *Encyclopedia of Applied Linguistics*, Malden, MA; Oxford: Wiley-Blackwell, 346-354.

Webb, S. (2019a). Introduction. In S. Webb, S. (ed.), *The Routledge Handbook of Vocabulary Studies.* London; New York: Routledge, 1-12.

Webb, S. (2019b). Incidental vocabulary learning. In Webb, S. (ed.), *The Routledge handbook of vocabulary knowledge*. London; New York: Routledge, 225-239

Webb, S. & Chang, A. (2012). Second language vocabulary growth. *RELC journal*, 43(1), 113-126.

Webb, S., Newton, J. & Chang, A. (2013). Incidental learning of collocation. *Language learning*, 63(1), 91-120.

Webb, S. & Nation, I. S. P. (2017). *How vocabulary is learned?* Oxford: Oxford University Press.

Wingate, U. (2004). Dictionary use - the need to teach strategies. *The Language Learning Journal,* 29(1), 5-11.

Wolfer, S., Cursio, M. N., Dias, M. S., Müller-Spitzer, C., & Domínguez Vázquez, M. J. (2018). Combining quantitative and qualitative methods in a study on dictionary use. In J. Čibej, V. Gorjanc, I. Kosem, & S. Krek (eds.), *Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts*. Ljubljana: Ljubljana University Press, Faculty of Arts, 101-112.

Wood, D. (2010). *Formulaic Language and Second Language Speech Fluency*: *Background, Evidence and Classroom Applications.* London; New York: Continuum.

Wood, D. (2015). *Fundamentals of formulaic language: An introduction*. London; New York: Bloomsbury.

Wood, D. (2019). Classifying and identifying formulaic language. In S. Webb (ed.), *The Routledge Handbook of Vocabulary Studies*. London; New York: Routledge, 30-45.

Wray, A. (2002). *Formulaic language and the lexicon*. New York: Cambridge University Press.

Wray, A., Bell, H. & Jones, K. (2016). How native and non-native speakers of English interpret unfamiliar formulaic sequences from Georgette Heyer's novel. *European Journal of English Studies*, 20(1), 47-63.

Yamamoto, Y. (2014). Multidimensional vocabulary acquisition through deliberate vocabulary list learning. *System* 42, 232-243.

Yanagisawa, A. & Webb, S. (2019). Measuring depth of vocabulary knowledge. In S. Webb (ed.), *The Routledge Handbook of Vocabulary Studies*. London; New York: Routledge, 371-386

Yongwei, G. (2012). Online English Dictionaries: Friend or Foe. *Proceedings of Euralex 2012*, 422-433.

Yoon, C. (2016). Concordancers and dictionaries as problem-solving tools for ESL academic writing. *Language Learning & Technology*, 20(1), 209-229.

Zeeland, H. & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening. *System*, 41(3), 609-624.

Zahar, R., Cobb, T. & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *Canadian modern language review*, 57(4), 541-572.

Zhang, L. J. & Zhang, D. (2019). Think-aloud protocols. In J. McKinley & R. Heath *The Routledge Handbook of Research Methods in Applied Linguistics*. London: Routledge, 302-311.

Zhao, N. (2009). Metacognitive strategy training and vocabulary learning of Chinese college students. *English Language Teaching*, 2(4), 123-129.

13. APPENDICES

APPENDIX 1: Task instructions in Finnish

Tutkimuksen ohjeistus (tutkijalle)

1. Kerro lyhyesti tutkimuksesta
    - sanastonoppimistehtävä max. 40min (ohjeistus kohta)
    - tehtävää tehdessä ajattelu ääneen
    - toiminta tietokoneella tallennetaan ruutukaappausvideolla, joka äänittää myös puheen
    - ruutukaappausvideo siirretään muistitikulle koneelta
    - vain minä katson/kuuntelen sitä
    - raportista ei selviä nimi, henkilöllisyys, mitä lukiota käy ym.
    - kun raportti on valmis, ruutukaappausvideo ja sen tekstiksi kirjoitettu versio tuhotaan
    - näistä lisää tietosuojailmoituksessa, jonka lähetän sähköpostilla
    - painota vapaaehtoisuutta, mahdollisuutta keskeyttää

2. Suostumuslomakkeen täyttö
    - versio itselle: tiedot + suostumus
    - versio tutkittavalle: pelkkä suostumus

3. Ääneen ajattelun harjoitustehtävä
    - seuraavaksi tehdään lyhyt tehtävä, jotta saisit kiinni ääneen ajattelusta
    - annan sinulle kohta kuvan ja tehtävänäsi on puhua ääneen ajatuksia, mitä kuvasta tulee
    - harjoitustehtävän ja myös oikean tehtävän aikana saatan kannustaa sinua puhumaan ääneen esimerkiksi sanomalla:”Mitä ajattelet nyt?”
    - Anna kuva, seuraa ja anna tarvittaessa palautetta/kehota puhumaan
    - Lopuksi: myös oikean tehtävän aikana ajattele samalla tavalla ääneen

4. Idiomien valinta
    - Sanastotehtävää varten etsitään ensin fraaseja, jotka on sinulle entuudestaan tuntemattomia
    - Jos tiedät joko osittain tai kokonaan fraasin merkityksen, sano että osaat fraasin.
    - Vaikka tunnistaisit jokaisen yksittäisen sanan merkityksen, pohdi tiedätkö mitä sanat tarkoittavat yhdessä
    - 7 ruotsinkielistä fraasia
    - 7 englanninkielistä fraasia

5. Ohjeistus:
    a. Edessäsi on nyt 7 englanninkielistä fraasia ja 7 ruotsinkielistä fraasia. Tavoitteena on, että selvität niiden merkityksen, eli mitä ne tarkoittavat. Kerron selkeästi, kun on aika vaihtaa kieltä. Voit aloittaa kielen X fraaseista. Sillä, ehditkö selvittää kaikkien 7 fraasin merkityksen ei ole väliä, sillä tarkoitus ole olla mahdollisimman nopea.
    b. Fraasien merkityksen selvittämisessä saat käyttää apuna näitä pöydällä olevia välineitä: tietokonetta, paperia, kyniä ja mitä tahansa muita keinoja, joista on sinulle hyötyä. Saat myös käyttää omaa puhelinta.
    c. Kun omasta mielestä tiedät mitä kyseinen fraasi tarkoittaa, kirjoita se lapulle
    d. Näytä esimerkki: mitä lappuihin kirjoitetaan (fraasi: kohdefraasi, sen merkitys).
    e. Lausetta voi halutessa käyttää apuna, mutta sitä ei tarvitse kirjoittaa lapulle.
    f. Tehtävää suorittaessa sano ääneen kaikki mitä luet, ajattelet ja millaisia huomioita teet, vaikka ne tuntuisivat itselle merkityksettömiltä. Saatan tutkimuksen aikana rohkaista sinua puhumaan. Muuten istun tuolla ja sinun ei tarvitse sen kummemmin välittää minusta.
    g. Kun olet kirjoittanut lappuun kyseisen fraasin merkityksen voit siirtyä seuraavaan fraasiin.
    h. Ei tarvitse ottaa paineita suoriutumisesta vaan teet mitä normaalistikin tekisit.

6. Onko kysymyksiä tutkimukseen liittyen tai jotain mikä tuntuu epäselvältä?

7. Aloitus: Sano ääneen, että nyt ruutukaappausvideo äänittää puheen ja tallentaa tietokoneella tapahtuvan toiminnan, jos sellaista on.

8. Kirjaa aloitusaika

9. Varmista, että tutkimushenkilö kirjoittaa lappuun oikeat asiat

10. Toteutus: Kannusta puhumaan ja kysy kysymyksiä, esim:
    - Kerro mitä luet/ ajattelet.
    - Älä unohda puhua/ lukea ääneen.
    - Mitä mietit nyt?

11. Seuraa aikaa
    - n. 20 min aikaa / kieli, huolimatta siitä montako fraasia on ehtinyt käydä läpi → yhteensä 40 min.

12. (Lisäkysymykset): Kysy mahdollisia lisäkysymyksiä tehtävän loputtua, jos jotain kiinnostavaa ilmenee.

13. Lopetus: Sano ääneen: “Nyt lopetan ruutukaappauksen tallennuksen”. Kiitä osallistumisesta ja anna tilaa keskustelulle tutkimuksesta, fiiliksestä jne.

14. Kerro, että tietosuojailmoitus tulee sähköpostissa. Painota osallistumisen perumisen mahdollisuutta ja vapaaehtoisuutta.

APPENDIX 2: Task instructions in English

Task instructions (for the researcher)

1. Tell briefly about the study
    - vocabulary learning task max. 40min (instructions will be given soon)
    - thinking aloud while completing the task
    - actions on the laptop will be recorded with a screen capture video that also records speech
    - video screen capture recording will be transferred from the laptop on a memory stick
    - only I will have access to the video screen capture recording
    - name, identity, and other personal information will not be revealed from the research report
    - when the report is ready, the screen capture recording and its transcription will be destroyed
    - more information on these in the data privacy announcement which will be sent via email
    - emphasize the voluntary nature of the study and the possibility to quit participating in the study

2. Filling in the consent form
    - my version: background information + consent form
    - the participants' version: only consent form

3. Think-aloud practice task
    - next we will do a short task so that you would familiarize with thinking aloud
    - I will soon give you a picture and your job is to think aloud any thoughts evoked by the picture.
    - during the practice task and the actual task, I may encourage you to talk out loud for example by saying: "What are you thinking now?"

- Hand out the picture, monitor and when necessary, give feedback/ encourage talking
- Finally: instruct the participant to similarly think aloud while completing the actual task.

4. Choosing idioms
    - For the vocabulary learning task we will now select phrases that are unfamiliar to you
    - If you know the phrase partially or fully, say that you know the phrase
    - Even if you would recognize the meaning of each single word, ponder whether you know what the words mean together.
    - 7 English phrases
    - 7 Swedish phrases

5. Task instructions:
    a. Now you have 7 English and 7 Swedish phrases in front of you. Your goal is to discover their meaning. I will tell clearly when it's time to change the language. You can start with the X (language) phrases. It does not matter whether you manage to discover the meaning of all 7 phrases because the goal is not to be as fast as possible.
    b. In discovering the meaning of these phrases, you can use items on the table as help: the laptop, paper, pencils and any other means that you find useful. You are also allowed to use your phone.
    c. When you think that you know what the phrase in question means, write the meaning on the answer sheet.
    d. Show an example of what the participant should write on the answer sheet (phrase + meaning)
    e. You can use the sentence on the card as help, but you are not supposed to write in on the answer sheet.
    f. While you are completing the task say everything that you read, think and any remarks that you make aloud even if they would feel insignificant to you. I may encourage you to speak during the task performance. Otherwise, I'll be sitting here and you do not need to pay attention to me.

g. When you've written the meaning of the phrase in question on the answer sheet, you can move on to the next phrase.
h. You don't need to take any pressure of your performance, just do what you normally would.

6. Do you have any questions regarding the study or anything that seems unclear?

7. Start: Mention that now the video screen capture is recording all speech and any on-screen activities.

8. Write the starting time

9. Ensure that the participant writes the right things on the answer sheet.

10. During the task: encourage the participant to speak and ask questions, such as:
    - Tell me what you are reading/thinking.
    - Don't forget to think/read out loud.
    - What are you thinking now?

11. Monitor time
    - approximately 20 minutes per language regardless of how many idioms have been discovered → altogether 40 minutes

12. (Additional questions): Ask possible additional questions if something interesting emerges.

13. Ending**:** Say out loud: "Now I'm stopping the video screen capture recording." Thank for participation and give room for discussion about the study, feelings etc.

14. Inform that the data privacy announcement will be sent via email. Emphasize the possibility of cancelling participation and the voluntary nature of participation.

APPENDIX 3: Background questionnaire and consent form for the researcher

Suostumus tutkimukseen osallistumisesta

Minua on pyydetty osallistumaan tutkimukseen englannin ja ruotsin sanaston oppimisesta. Olen saanut tarpeeksi tietoa tutkimuksesta. Noomi Halonen on myös vastannut kaikkiin kysymyksiini tutkimuksesta.

Ymmärrän, että tähän tutkimukseen osallistuminen on vapaaehtoista. Minulla on oikeus, milloin tahansa tutkimuksen aikana keskeyttää tutkimukseen osallistuminen. Minun ei tarvitse kertoa keskeyttämisen syytä, eikä siitä aiheudu minulle mitään ikäviä seuraamuksia.

Tietoja osallistujasta

☐ Olen lukion 2. vuoden opiskelija. / ☐ Olen lukion 3. vuoden opiskelija. / ☐ Olen lukion 4. vuoden opiskelija.

☐ Olen nainen. / ☐ Olen mies.

☐ Olen jo suorittanut englannin yo-kokeen. ☐ Olen jo suorittanut ruotsin yo-kokeen.

Aion kirjoittaa / Olen kirjoittanut ruotsin pitkänä (A-ruotsi) ☐ / keskipitkänä (B-ruotsi) ☐

Kirjoita oma arviosi lukioajan kurssien keskiarvosta ja omasta koetusta kielitaidosta arvosanalla 4-10:

Englannin kurssien keskiarvo n. _____ Englannin kielitaito: _____

Ruotsin kurssien keskiarvo n. _____ Ruotsin kielitaito: _____

Suostumus

☐ Haluan osallistua tutkimukseen.

☐ Tutkimustilanteen saa äänittää.

☐ Mahdollisen toiminnan tietokoneella saa tallentaa ruutukaappausvideolla.

☐ Ruutukaappausvideosta otettuja kuvia saa nimettömästi julkaista osana tutkimusraporttia.

_____

Päiväys

_____

Tutkittavan allekirjoitus

_____

Tutkittavan nimen selvennys

_____

Tutkijan allekirjoitus

_____

Tutkijan nimen selvennys

APPENDIX 4: Consent form for the participant

Suostumus

☐ Haluan osallistua tutkimukseen.

☐ Tutkimustilanteen saa äänittää.

☐ Mahdollisen toiminnan tietokoneella saa tallentaa ruutukaappausvideolla.

☐ Ruutukaappausvideosta otettuja kuvia saa nimettömänä julkaista osana tutkimusraporttia.

_____

Päiväys

_____

Tutkittavan allekirjoitus

_____

Tutkittavan nimen selvennys

_____

Tutkijan allekirjoitus

_____

Tutkijan nimen selvennys

APPENDIX 5: Data privacy announcement

Tietoa tutkimukseen osallistuvalle

Olen Noomi Halonen, Jyväskylän yliopiston opiskelija, ja tutkin oppimisstrategioita, joita käytetään englannin ja ruotsin idiomien merkitysten selvittämisessä.

Tässä tutkimuksessa käsittelen tutkimuksessa saamiani tietoja yleisen edun perusteella ja olen pyytänyt sinulta suostumuksen osallistua tutkimukseen.

Vapaaehtoisuus ja tutkittavan oikeudet

Tutkimukseen osallistuminen on täysin vapaaehtoista. Voit kieltäytyä, keskeyttää tai peruuttaa tutkimukseen osallistumisen milloin tahansa. Sinun ei tarvitse kertoa minulle, miksi et halua osallistua. Sinulla on oikeus tarkistaa tai oikaista antamasi tiedot, voit myös kertoa minulle, että et halua tietojasi käsiteltävän ja voit tehdä valituksen henkilötietojen käsittelystä.

Jos sinulla on kysyttävää oikeuksistasi tai muihin tutkimukseen liittyvistä asioista, voit olla yhteydessä joko minuun, Noomi Haloseen nosikurk@student.jyu.fi, p. 0408333286 tai tietosuojaan liittyvissä asioissa myös yliopiston tietosuojavastaavaan tietosuoja@jyu.fi, p. 040 805 3297.

Tietoa tutkimuksesta

Tutkimus suoritetaan sovitussa paikassa ja sovittuun aikaan. Tutkimus kestää kokonaisuudessaan n. tunnin. Tutkimus koostuu oppimistehtävästä, jossa tutkittava selvittää englannin- ja ruotsinkielisten idiomien merkityksen haluamillaan tavoilla. Tutkimustilanne äänitetään ja mahdollinen toiminta tietokoneella tallennetaan ruutukaappaus-videolla. Sanastontehtävän vastaukset kerätään talteen.

Aineiston säilyttäminen ja hävittäminen

Ruutukaappaus-video (sis. äänitteen) siirretään tutkimuksen jälkeen välittömästi muistitikulle. Ainoastaan tutkijalla on pääsy niihin ja oikeus katsoa/kuunnella niitä. Ruututoimintaa tallentavaa ruutukaappaus-videota voidaan hyödyntää tutkimustulosten esittelyssä esimerkiksi yksittäisinä kuvina näytöllä tapahtuvasta toiminnasta. Tutkimusraportin valmistuttua ruutukaappausvideo (sis. äänitteen), transkriptio (puhe tekstimuotoon kirjoitettuna + kuvaus tietokoneella tapahtuvasta toiminnasta) sekä sanastotehtävän vastaukset hävitetään.

Henkilötietojen käsittely

Käsittelen tutkimuksen aikana saadut tiedot luottamuksellisesti ja nimettömästi.

Tutkimustulosten raportissa ei mainita osallistujan nimeä, luokkaa, lukion nimeä, asuinpaikkaa tai muitakaan henkilötietoja, joiden perusteella tutkittavan henkilöllisyys voi paljastua. Ainoastaan kuvia ruutukaappausvideoista voidaan liittää raporttiin.

Tutkimusraportissa kerrotaan vain tutkimustilanteessa ilmenevistä tuloksista. Lisäksi tutkittavan luokka-aste (esim. lukion 2. vuoden opiskelija), sukupuoli, onko jo tehnyt yo-kokeen, arvio kurssiarvosanoista ja arvio omasta kielitaidosta voidaan mainita.

Tutkimuksen tulokset

Tutkimuksen tulokset julkaistaan pro-gradu tutkielmassani Jyväskylän yliopiston tutkimusarkisto JYX:issä. Tutkimuksen tulokset ovat kaikkien halukkaiden hyödynnettävissä.

Tutkittavan oikeudet

Voit kysyä minulta mitä tahansa tutkimuksesta ennen tutkimusta, sen aikana tai sen jälkeen. Sinulla on oikeus tarkastaa tai oikaista antamasi tiedot, voit myös kertoa minulle, että et halua tietojasi käsiteltävän tai tehdä valituksen henkilötietojesi käsittelystä.

APPENDIX 6: List of English idioms and example sentences

English idiom	Dictionary definition	Example sentence
hot air	If something that someone says or writes is hot air, it is not sincere.	All their talk about cooperation is just so much **hot air**. There are still endless disagreements original word *disputes* was changed into disagreements
a can of worms	a situation or subject that is very complicated, difficult or unpleasant to deal with or discuss	Now we have uncovered **a can of worms** in which there has not only been shameful abuse of power, but a failure of moral authority of the worst kind.
a red herring	If something is a red herring, it takes people's attention away from the main subject, problem, or situation that they should be considering.	A sighting of the missing woman in London turned out to be **a red herring**.
make the grade	you succeed at something, usually by reaching a particular standard	As a child, she wanted to be a dancer but failed to **make the grade**.
jump the gun	do something before the right time	The book wasn't due to be released until September 10 but some booksellers have **jumped the gun** and decided to sell it early.
make noises	If you make noises about something you might do, you mention it briefly in a way that is not definite or detailed.	His mother had started **making noises** about it being time for him to leave home.
play games	If someone plays games, they are dishonest about what they know or intend to do in order to gain an advantage.	"Don't **play games**, Mona. I know about the theft, and I know that you know about it, I repeat: why are you here?"
bend over backward/ backwards	If you bend over backwards, you try very hard to help or please someone, even though it causes you trouble.	We **bent over backwards** to make them feel welcome and they didn't thank us once.
come home to roost	If something bad that someone did comes homes to roost, it now causes problems for them.	You should have known that your lies would **come home to roost** in the end. original phrase *ought to* was changed into *should*

give someone a run for their money	If you give a very skillful person or team a run for their money in a competition, you compete as well as them, or almost as well.	We think the Irish will give the Welsh a good **run for their money**.
hold out an olive branch	you say or do something to show that you want to end a disagreement with them.	We are **holding out an olive branch**, inviting the landowners to talk to us.
let your hair down	If you let your hair down, you relax and enjoy yourself, and do not worry about being dignified or behaving correctly.	It is only with friends that most people feel they can **let their hair down** and be themselves.
take it on the chin	If you take it on the chin, you bravely accept criticism or a difficult situation.	When the police arrived, he **took it on the chin**, apologizing for the trouble he'd caused them.
cut no ice	If you say that something cuts no ice with you, you mean that you are not impressed or influenced by it.	Statistics **cut no ice** with anyone scared of going up in the air in a plane.
go out on a limb	If you go out on a limb, you do something or say something that is different from what most people do or say and is therefore risky.	He does not want to **go out on a limb** and try something completely new.
make a splash	If someone or something makes a splash, they attract a lot of attention, often by being very successful.	His debut single comes out in May - but has already **made** quite **a splash** in the States.
have it in for someone	If someone has it in for you, they do not like you and they want to make life difficult for you.	He has always **had it in for** the Dawkins family.
take the edge off something	If something takes the edge off a feeling or situation, especially an unpleasant one, it weakens its effect and unpleasantness.	My head never seemed to clear completely, and the painkillers only **took the edge off** my pain.
on the ropes	If a person or an organization is on the ropes, they are very close to failing or being defeated.	Far fewer tourists are coming to this country and as a consequence, our hotel industry is **on the ropes**.
in the loop	If someone is in the loop, they are part of a group of people who	The vice president was almost certainly **in the loop**.

	have information about a particular thing.	
from scratch	If you do something or start something from scratch, you create something completely new, rather than adding to something that already exists.	He would rather start again **from scratch** with new rules, new members, and a new voting system. original phrase *electoral system* was changed into *voting system.*
lock, stock and barrel	You use lock, stock and barrel to talk about every part of something	It would have been much easier for us to have shut the business down **lock, stock and barrel**.
in the groove	If someone, especially a sports person or team is in the groove, they are performing well.	Nick is **in the groove** as he showed with seven goals last weekend.

APPENDIX 7: List of Swedish idioms, example sentences and frequency

Swedish idiom	**Dictionary definition**	**Example sentence**	**Google hits** (frequency) **and date of googling**
en annan femma	någonting annat än det man tänkt sig; en annan sak	"Jaså, ingår också frukost och lunch i resans pris! Det var **en annan femma**. Då köper jag den."	(182 000) 17.2.2022
för egen maskin	på egen hand; av egen kraft	Efter att de hade fått motorstopp fick de ta sig hem **för egen maskin**.	(107 000), 17.2.2022
gå in i väggen	inte orka mer; bli överansträngd eller utbränd	Det hade varit alldeles för jobbigt och stressigt på arbetsplatsen det senaste året. Flera anställda hade **gått in i väggen** och var därför sjukskrivna.	(339 000), 10.3.2022
ge svar på tal	svara någon på ett sätt så att han/ hon blir tyst	När hon kallade henne för idiot, **gav** hon honom **svar på tal**. original word *kärring* was changed into *idiot*	(34 900), 10.3.2022
gå i taket	bli väldigt arg eller upprörd; explodera av ilska	Jag höll på att **gå i taket** när jag blev anklagad för något jag inte hade gjort.	(243 000); 10.3.2022
äga rum	hända; kommer att inträffa	Rättegången kommer att **äga rum** onsdagen den 3 mars.	(2 370 000), 10.3.2022
vara med på ett hörn	vara med lite vid sidan av	Även om jag inte är så bra på att spela teater kanske jag ändå kunde **vara med på ett hörn**.	(142 000), 10.3.2022
vara hemma på något	vara duktig på något	Han är **hemma på** datorer.	(50 600), 10.3. 2022
mellan varven	före eller efter då något händer; en period när det är lugnare	**Mellan varven** när vi har mindre att göra kan vi ta det mycket lugnare.	(405 000), 10.3.2022
vända på steken	tänka eller göra på ett annorlunda sätt; se	"Den här talet kan jag inte lösa hur mycket jag än försöker. Det går nog lättare om man **vänder på**	(66 500), 10.3.2022

	något från en annan synvinkel	**steken** och börjar med att dividera.”	
när det kommer till kritan	när det verkligen gäller, i ett avgörande läge	**När det kommer till kritan**, brukar han hjälpa oss.	(90 100), 17.2.2022
dött lopp	en tävling där två kommer i mål samtidigt; alldeles jämnt.	Det blev **dött lopp** mellan båda löparna.	(79 000), 10.3.2022
sätta ner foten	säga ifrån; protestera mot någonting	Det gäller att våga **sätta ner foten** när man tycker att något är helt fel.	(122 000), 17.2.2022
göra en djupdykning	studera något i detalj	Vi behöver nog **göra en djupdykning** och granska det hela närmare.	(98 700), 10.3.2022
ett plus i kanten	extra bra, få beröm	Att han har arbetat utomlands är **ett plus i kanten**.	(130 000), 10.3.2022
ta itu med något	1) börja med något som känns tråkigt, 2) reda upp något med någon	Nu måste jag **ta itu med** trädgården.	(1 380 000), 10.3.2022
se om sitt hus	tänka på sig själv och planera så bra som möjligt	Det gäller nog att **se om sitt hus** och köpa biljetter i tid innan de är slutsålda.	(44 800), 10.3.2022
sitta lugnt i båten	man avvaktar och gör inget	Trots att aktierna går ner fick han rådet att **sitta lugnt i båten**.	(23 700), 10.3.2022
sticka ut hakan	våga säga sin uppriktiga mening även om det ogillas	Han är inte rädd för att **sticka ut hakan** även om han får en del kritik efteråt. original word *obehag* was changed into *kritik*	(61 700), 10.3.2022
peka med helan handen	vara extra tydlig med att något måste göras omgående	Chefen var irriterad över att inget blev gjort och **pekade med hela handen** om vad som måste göras.	(49 700), 17.2.2022
till äventyrs	kanske; möjligen	Har du t**ill äventyrs** en penna att låna mig?	(96 200), 17.2.2022

göra rätt för sig	inte leva på andra; inte ligga någon till last	Han ville **göra rätt för sig** hemma och betalade för mat och husrum.	(118 000), 10.3.2022
skynda långsamt	ta det lugnt och inte stressa	Det är bättre ibland att **skynda långsamt** när man ska göra någonting svårt.	(97 000), 20.3.2021

APPENDIX 8: English translations of the Finnish quotations

Quotation 1

Original excerpt (1):
[…] sticka ut hakan, vois olla johonki virheitten tekemiseen liittyvä, se ei oo, niin, se ei oo peloissaan, vaikka kuinka sais huonoa palautetta jälkeenpäin. sticka ut hakan, sticka ut, kuulostaa niinku vähän semmoselta, mikä se on niinku, öö ampua ohi. (Participant E)

English translation (1):
[…] *sticka ut hakan*, could be something related to making mistakes, it is not, well, it (he) is not afraid despite receiving negative feedback afterwards. *sticka ut hakan*, *sticka ut*, it sounds kind of like, what is it like, umm to miss a target. (Participant E)

Quotation 2

Original excerpt (2):
"Tilastot eivät riko jäätä [unintelligible]. pelkää. ilmaa lentokoneessa." […] "Tilastot eivät riko jäätä, kenenkään kanssa joka pelkää mennä ylös ilmaan lentokoneessa." […] "Tilastoilla ei ole vaikutusta, eli tilastoilla ei ole mitään vaikutusta siihen, että, tilastot ei vaikuta siihen että paljonko ihmiset pelkää lentämistä." (Participant G)

English translation (2):
"Statistics cut no ice. [unintelligible]. afraid. air airplane." […] "Statistics cut no ice, with anyone who is afraid of going up in the air in an airplane." […] "Statistics have no influence, so statistics have no influence on that, statistics have no influence on how much people are afraid of flying. (Participant G)

Quotation 3

Original excerpt (3):

“Ta itu med, mikäs tuo trädgården on, voiko sitä ottaa paloiks?” […] “Nyt minun täytyy ottaa palasiksi puutarhan kanssa. öö, vai onks se niinku puutarhassa yleensä tehdään jotain joten se vois olla ryhtyä tekemään jotain.” (Participant A)

English translation (3):
“*Ta itu med*, what is that *trädgården*, can it be taken into pieces? […] “Now I need to take into pieces with the garden. umm, or is it like in the garden you usually do something so it could be undertake doing something.” (Participant A)