UNIVERSITY OF TURKU

Idioms in L2 English Essays by Swedish University Students

Idiom Types and The Effect of Spending Time in an English-speaking Environment

Tomi Vuorinen

Master's Thesis

Degree Programme for Language Learning and Teaching, Department of English

School of Languages and Translation Studies

Faculty of Humanities

University of Turku

April 2022

Master's Thesis

Degree Programme for Language Learning and Teaching, Department of English:
Tomi Vuorinen
Idioms in L2 English Essays by Swedish University Students: Idiom Types and The Effect of Spending Time in an English-speaking Environment
Number of pages: 41 pages, 1 appendix

The topic of the thesis is English idioms in essays written by Swedish university students and the effect of spending time in an English-speaking environment has on their idiom usage. The material of the study is collected from The Uppsala Student English Corpus (USE) which is available online. 50 texts were collected from the corpus: 25 from students with 0 to 6 months of experience and 25 from students with more than 6 months of experience. The essays were read and checked for idiomatic expressions while also marking down the form of the idiom. The three forms of idioms were literal, semi-literal and pure idioms, based on idiom classification by Fernando (1996). The study is exploratory in a quantitative manner.

The main results are that the group with over 6 months of experience used more idioms, but when examining the relationship between the amount of experience reported as months, no link between higher numbers of months of experience and more idiom usage was found. In fact, more months of experience seemed to lead to less idiom usage. The conclusion is that over 6 months of experience in English-speaking environment has a positive effect on idiom use but there might be a threshold after which the effect diminishes. The suggestion is that this material should be examined more closely with quantitative methods using a higher number of essays.

Table of contents

1   Introduction

Do you know what is meant by the expression *a piece of cake*? Do you know that it has nothing to do with cake at all? These types of expressions with non-literal meanings are called idioms and they are found in various kinds of texts, but mainly in everyday spoken language and informal texts. They make the language more vivid and interesting while presenting ideas and opinions about different subjects in a compact manner. When we speak, we do not look always look for single words, but lexical chunks consisting of multiple words. Think about greetings like *How do you do?* They seem to come out of your mouth as a continuous string of words without the need to stop and think about every single word choice in the sentence. It is easy to pay no attention to such lexical chunks as they are used so automatically. But idioms and other lexical chunks have been a field of research for many years now even though their significance was not always recognized. Also, consider the fact that lexical chunks seem to be fixed, meaning that you cannot really do any alterations to them. How does *blue deep sea* sound to you compared to *deep blue sea*?

There is no one universally accepted definition for an idiom. Instead, there are multiple competing definitions. Most definitions seem to agree on the criteria of idioms being multi-worded non-literal expressions with some degree of fixedness or frozenness as other researchers call it. Idioms are not easy for second language learners as idioms differ between languages: *kill two birds with one stone* and *kaksi kärpästä yhdellä iskulla* (two flies with one swat) have the same meaning but the exact words are different. This can understandably lead to confusion. However, second language idioms are important to learn for understanding various kinds of texts correctly and being able to use idioms in a foreign language makes the learner seem more native-like. A misunderstood idiom can lead to awkward situations. According to the Common European Framework of Reference for Languages (CEFR), to reach a high level of spoken language skill (C2 level), one must have "good command of idiomatic expressions and colloquialisms" (Council of Europe 2022).

It has been proposed that a good method for learning second language vocabulary, which idioms are a part of, is incidental learning where, for example, the learner spends time and learns to communicate in a second language environment in a natural way. The aim of this study is to examine how an advanced student's experience in an English-speaking environment affects the frequency of idioms used. The aim is to find out whether more experienced students use more idiomatic language and to find out how they use it. Are they

using very simple idioms with easily understood meanings or more difficult ones with a meaning that is more non-literal? The study is exploratory in the sense that the method is quantitative, but the limited data does not allow for more specific statistical analysis. The aim is to find a link between the experience and number of idioms used and find out whether there is a tendency for more experienced students to use more idiomatic language.

First, I will present and discuss the theoretical background of this study, starting from studies related to second language vocabulary acquisition and then moving onto the topic of idioms. Then, I will discuss the details of the present study, including the research questions, the material, and the methods of analysis. Then, I will analyze the data and discuss the results of the analysis. After that, I will end the thesis with the conclusions of the study and give ideas for the purposes of carrying out future research about this topic.

2   Theoretical background

As idioms are a part of the wider area of vocabulary, this chapter will first present some previous studies on second language vocabulary learning and then discuss studies about formulaic language which idioms are also a part of. Then the various definitions of idioms will be presented and finally there will be information about idioms in second language acquisition along with previous studies on it.

2.1   Second language vocabulary learning

Before, vocabulary was seen as secondary to grammatical structures in the field of language learning and teaching, but vocabulary's importance is now better recognized (Mäntylä 2004, 17). It was not until the 1970s and 1980s, relatively late, that researchers started to pay attention to teaching and learning vocabulary (Niitemaa 2014, 138). Also, the difficulty of learning vocabulary has been recognized:

> Achieving native-like command of second language vocabulary poses a real challenge. It may well be easier to master a system of rules, such as the grammar of a language, than an ever-growing class of lexical items. What is more, it is not merely the size of the L2 mental lexicon that matters, but also the appropriate use of the words that one has access to. (Pietilä et al 2015, 1)

It is not truly known what happens when words are learned (Niitemaa 2014, 147). In the field of vocabulary studies, it is known that knowing a word consists of two parts: reception and production (ibid., 143). Reception means to recognize a word and production refers to being able to use the word in speech, for example (ibid.) Receptive vocabulary of a learner is usually larger than the productive one. Language transfer, including lexical transfer, started to gain importance in the 1960s with the occurrence of contrastive analysis (Jajić Novogradec 2021, 1). Language transfer refers to the effect of L1 on L2 language learning and it can be both positive and negative. The importance of learning chunks of words such as idioms and collocations has also been noted (Niitemaa 2014, 141).

Vocabulary knowledge can be assessed in various manners (Mäntylä 2004, 18). One way is through reading comprehension: examining how many words learners know in a foreign language and how many they should be familiar with to be able to understand a certain text (Niitemaa 2014, 142). The words known by a speaker can be referred to as the *mental lexicon* (ibid., 145). One estimate is that a learner should know about 3000 word families to be able to understand about 95 percent of English language texts (ibid.) The modern thinking is that one

should be able to understand 98 percent of English texts which would mean knowing about 8000 word families but knowing 2000 word families, the learner can get pretty far (ibid.). Word family refers to the base word and all the possible forms along with derivates (ibid., 140). Estimates like this are usually based on frequency lists that are available as computerized corpora in the modern times. "Frequency lists have led to the search for core vocabulary", which is controversial as the needs of learners vary (Mäntylä 2004, 18). Core vocabulary is not likely to include metaphorical expressions such as idioms (ibid.) even though they are seen as important in the modern age as Niitemaa (2014, 141) stated.

L1 can enhance the learning of a foreign language learning due to pre-existing general knowledge about how language works (Mäntylä 2004, 19). The effect of L1 can also be negative. A great example are false friends: a word appears to have a very similar equivalent in the foreign language, but it actually has a different meaning in the L2 (ibid.) Swan (1997, 3-6) gives a list of issues that can occur as a result of cross-linguistic transfer. A language learner can, for example, borrow directly from his mother tongue or apply L1 semantics to an L2 word, for example, say *open the TV* in English if Finnish is their L1 (Mäntylä 2004, 19-20). Idioms can be found in all languages and are formed within similar guidelines (ibid., 20). Differences can be noted when the semantics and literal meanings of an idiom are taken into closer examination (ibid.) Ringbom (1987, 135) makes the claim that "the learner tends to assume that the system of L2 is more or less the same in his L1 until he has discovered it is not." According to Mäntylä (2004, 20), "With less frequent expressions and more specialised language, this discovery may come at a fairly late stage, if at all."

L2 teaching used to focus on grammar whereas it now focuses on vocabulary (Mäntylä 2004, 20). The learner needs to have a wide vocabulary in order to use grammar and apply the rules of it (ibid.) Teaching foreign language vocabulary has changed over the years, too (ibid.) Language teaching has recognized the importance of thoroughly analyzing what knowing a word means and the stages involved in the learning process (ibid., 20-21). Mäntylä also states the following:

> In addition to reconsidering the role of vocabulary in language and the properties of vocabulary, vocabulary teaching has also been affected by what have been considered to efficient language teaching methods and most useful goals in language learning. Different approaches have been valued at different times and vocabulary has likewise been taught by different methods each considered valuable in its time for the desired purpose. (Mäntylä 2004, 21)

A well-known early method of teaching vocabulary is the grammar-translation method which was introduced in Prussia at the end of the 1700s (Zimmerman 1997):

> Students were provided detailed explanations of grammar in their native languages, paradigms to memorize, and bilingual vocabulary lists to learn; these prepared them for the regular task of translating long passages of the classics. (Zimmerman 1997, 6)

According to Järvinen (2014, 92), the grammar-translation method focused on the right form of language and the students were evaluated on the basis of the grammatical errors they made. The content of the texts used in the method did not matter that much (ibid.) Most of the classes were similar and routine-like (ibid.) and the lack of variation may cause boredom in the students which probably the reason for the method being less popular in the modern times. The words taught in schoolbooks are not chosen at random; they are based on frequency lists and belong to a certain theme the chapter of the book is presenting (Mäntylä 2004, 21). Using the L1 in definitions and explanations is advantageous as studies "have shown that for many learners learning is faster if the meaning of a word is conveyed by a first language translation." (Nation 2001, 66). On the other hand, relying on L1 is not free of issues (Mäntylä 2004, 21). When L1 is commonly used in the classroom, the learners do not get used to using the L2 the same way they would if the classes were in L2 (Nagy 1997, 73-74).

A natural approach, also referred to as incidental vocabulary learning relies on the use of L2: learners can attempt to guess meanings with the help of the clues provided by the context or by using L2 dictionaries (Sökmen 1997, 237). Incidental encounters with words, often slow and prone to errors, are just one method of facilitating vocabulary acquisition (ibid., 237-239). Repetition and opportunities for using the newly learnt word are also very important. Baroto (2017, 2) claims that "Input is a central aspect in learning second language. Increased input, or exposure, of the target language will contribute to lasting better English skills." However, exposure is not enough for drastically expanding a learner's vocabulary (Vaurio 1998, 47). Waldemar Marton's 1977 study "Foreign Vocabulary Learning as Problem No. 1 of Language Teaching at the Advanced Level" shows that even a long visit in an English-speaking country does not expand the learner's vocabulary much although an increase in receptive vocabulary has been examined in advanced learners (Marton in Mäntylä 2004, 21-22). What comes to L2 phrasal language, exposure is not sufficient to increase even an advanced learners' knowledge of it (ibid., 22). "More and more research points to the ineffectiveness of just using implicit vocabulary instruction" (Sökmen 1997, 238). However, reading longer texts, when it is well

guided, appears to have the potential for improving aspects of spelling, meaning and grammar when acquiring new vocabulary (Coady 1997, 235). Incidental vocabulary acquisition is linked to language proficiency, and they mutually reinforce each other (Vaurio 1998, 47). This is closely related to the purpose of the present study as the objective is to examine the effect of time spent in an English-speaking environment on idiom usage by learners.

"Topics and themes have become increasingly important in vocabulary teaching as the significance of word relations and connecting language and content have been acknowledged" (Mäntylä 2004, 22). Rather than just individual words, vocabulary should be seen as a network of words and meanings relating to each other (ibid.) Words are networked in an associative manner in the mental lexicon (Niitemaa 2014, 145). The idea has been further developed into teaching content in L2: the method emphasizes using the L2 without using the L1 as a mediator (Mäntylä 2004, 22). It has the advantage of requiring the learner to focus on meaning rather than on form (Brinton et al. 1989, 4) as they need to attempt linking the old information to the new information. The reading method and situational language teaching are connected to the idea of working only in L2 (Zimmerman 1997, 9-10). According to Krantz (1991, 10), reading is considered a major means of increasing vocabulary in moderately advanced L2 students. The method of reading involves much more than just reading as the learner must first notice the word, decontextualize it and by negotiation, or some other way comprehend its meaning (Nation 2001, 63-65). Nation (ibid., 85) lists some ways to communicate the meaning of a word: by performing actions, by showing objects, showing pictures or diagrams or giving an L1 or L2 definition, for example. The choice for the method of communicating the meaning is based on two considerations: the reason for explaining the meaning of it and the degree to which the method of explaining it represents the wanted meaning for it (ibid.)

Meaning has become a significant target of learning as the importance of a wide vocabulary has been noted as an important goal of language teaching (Mäntylä 2004, 23). This is evident in different communicative approaches where the main focus is on language use and functions instead of forms (Zimmerman 1997, 12-13). Unlike grammar-translation method, the focus is not on reading and translating but on spoken language and oral fluency (Mäntylä 2004, 23). Good learners can pay attention to both meaning and form and they can shift their focus between these two areas (Ellis 1994, 546-549), as is required by this method (Mäntylä 2004, 23). Ellis (1994, 556) also claims that "social and interactional strategies may be more important with younger learners". However, Mäntylä (2004, 23) disagrees with this view:

> this may not always be true as different people respond to different methods in different ways, and also the tasks and goals of learning set demands for methods. Nevertheless, with younger learners, interactive and communicative approaches may prove more fruitful than traditional classroom instruction. (Mäntylä 2004, 23)

When it comes to vocabulary acquisition, there is little evidence for the effectiveness of communicative approaches (Mäntylä 2004, 23). They do, however, guide the learners to actively use language and they can enhance vocabulary acquisition too (ibid.) Furthermore, extensive vocabulary is useless without knowing how to use it, so form is also important (ibid.) The communicative methods may not always focus enough on the levels of knowing a word, its interpretations, restrictions in use, metaphoricity and other factors (ibid.).

Vocabulary teaching has changed according to the current trends in language teaching in general, the areas given importance in language and how vocabulary has been perceived (Mäntylä 2004, 23). It appears that these days the variety of methods and strategies is recognized along with the learners' personal differences and varying goals for learning (ibid.) The complexity of vocabulary research is reflected in research as the number of studies on vocabulary acquisition is rising. "In language learning and teaching, particularly in SLA and vocabulary teaching the emphasis has in recent years been placed not just on single words but larger entities, "chunks"" (ibid., 24). Cowie (1998, 1) states that "phraseology has now become the major field of pure and applied research for Western linguists that it had, much earlier, for scholars in the former Soviet Union and other countries of Eastern Europe."

2.2 Formulaic language

Now that I have discussed the field of second language vocabulary acquisition, the next topic is formulaic language which is a crucial part of vocabulary and closely related to idioms and idiomaticity. Discourse analysis, the analysis of language beyond the syntactic level, experienced a strong revival in the 1970s and 1980s (Wulff 2008, 10). The major accomplishment of various discourse analysis approaches with regards to idiomaticity is that they emphasize that phraseological units are not a marginal phenomenon in language (ibid.). In fact, they are highly prominent and indispensable units of language (ibid.). In the field of phraseology, there is no established terminology to refer to these units (Mäntylä 2004, 24). Moon (1998, 2), for example, talks about fixed expressions to "cover several kinds of phrasal lexeme, phraseological unit, or multi-word lexical item" and Wulff (2008, 10) talks about phraseological units. Prefabricated units or patterns, and formulaic language are also frequently used for pieces of language over one word, such as collocations, phrases, sayings,

greetings, and idioms (Mäntylä 2004, 24). Wray and Perkins (2000) propose that formulaic language is:

> a sequence, continuous or discontinuous, of words or other meaning 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 (Wray and Perkins 2000, 1)

In phraseology, rather than assigning idioms a sort of special status like early generative approaches did (Katz and Postal 1964), idioms are considered a subcategory of multi-word units.

Formulaic language forms a bridge between grammar and vocabulary. *How are you?* is a question that can be analyzed grammatically, but it is also a unit that learners learn to use without analyzing the separate components of the structure (Nattinger and DeCarrico 1992, 11). The question is used in a particular context. The learner also learns how to respond to this question as a prefabricated single unit rather than as two separate sentences: *I'm fine thanks. How are you?* (Mäntylä 2004, 24). Regina Weinert, in her 1995 review titled "The Role of Formulaic Language in Second Language Acquisition: A Review", published in *Applied Linguistics* 16 (issue 2) refers to multiple studies focusing on multi-word or multi-form strings (such as *can't*) which are produced or recalled as a chunk instead of as multiple separate words (Weinert in Mäntylä 2004, 24). They are more alike to single words rather than structures of multiple words formed according to linguistic rules.

Cowie (1998, 1) argues that reaching native-level proficiency "depends crucially on a stock of prefabricated units". One must remember, though, that not all learners aim to become native-level and it is not necessary (Mäntylä 2004, 25). Knowledge of formulaic language aids comprehension: it is crucial to know that *kicking the bucket* has nothing to do with either actual kicking or a bucket in order to comprehend the expression (ibid.) Similarly, not knowing the purpose and function of frequently used expressions like *you know* prevents "understanding and interpreting the text and situation" (ibid.) Even beginner and intermediate learners need phraseological knowledge for comprehending language correctly (ibid.) Repetition and imitation also have their place in language learning, and frequently used patterns can be easier to memorize (ibid.)

I have now discussed formulaic language and its importance in achieving a high level of language skill. Next, I will discuss idiomaticity and idioms which are both closely related to

formulaic language. I will focus on the various definitions of an idiom and pick one definition to use for the purposes of this study.

2.3   Definition and the characteristics of an idiom

In general, idioms are not well-defined and various definitions exists. It is important to note that the terms idiomaticity and idiom are closely related but not identical by any means. Oxford Learner's Dictionaries (2022, s.v. idiom) defines an idiom as "a group of words whose meaning is different from the meanings of the individual words" and gives *let the cat out of the bag* as an example of one. In fact, it is only during the 20th century that the term idiom has come to mean what it means today (Nenonen 2002, 2-3). Before that, the term only referred to a certain style of speaking or language of some group. Idioms are similar to other types of lexical chunks such as collocations, phrases and proverbs (Mäntylä 2006, 152) and can be difficult to separate from one another. They all seem to be learned as chunks instead of single words and they are all used as fixed expressions. Fixed expression is a term "used to cover several kinds of phrasal lexemes, phraseological units, or multi-word lexical items" (Moon 1998, 2). Mäntylä (2006, 153) defines an idiom as a figurative expression with two or more words where the true meaning of the expression differs from the literal meaning. This feature of an idiom is referred to as "opaqueness". Mäntylä (2004, 28-35) lists several characteristics of idioms: metaphoricity, analysability/non-compositionality, fixedness of form, level of formality and multi-wordedness. Multi-wordedness as an essential feature of idioms is not without problems, though (Mäntylä 2004, 35). Mäntylä (ibid.) gives the multi-worded non-literal expression *How do you do?* as an example of a case where idiomaticity is under questioning regardless of the expression being multi-worded.

Wood (1986, 95) proposes the following definition of idioms: "An idiom is a complex expression which is wholly non-compositional in meaning and wholly non-productive in form". Häkkinen (2002, 451) and Nenonen (2002, 7-8) state that idioms are conventional, which means that an idiom is familiar to the members of a language community and is in common use. Previously known idioms can aid in understanding an idiom that is at first unrecognizable to the speaker and thus the idiom can remain conventional even if it seems unfamiliar at first sight. This shows that idioms are tied to certain cultures: an idiom used in another culture might not be understandable in another one.

As is the case with idioms, there are major debates and multiple proposals on the definition of idiomaticity. As stated earlier, idioms are closely related to idiomaticity, but the terms do

not refer to the same thing. For example, habitual collocations such as *rosy cheeks* show idiomaticity while not being idioms due to reasons such as being relatively unrestricted in adjectival and nominal variants (Fernando 1996, 30). Idiomaticity refers to a phenomenon that is multifactorial in nature, meaning that it comprises of more than compositionality alone (Wulff 2008, 8). Idiomaticity is the feature of phrases taking on meanings that go beyond the literal meanings of the words contained in the phrase (ibid.). Idiomaticity is the opposite of compositionality. Idiomaticity is scalar in nature in the sense that different constructions in language can be idiomatic on different levels, creating a continuum (ibid.). It is also a feature that cannot be regarded as marginal, but rather deserving of a central position in any grammatical theory (ibid.). Fernando and Flavell (1981) give the following five-part criteria for idiomaticity:

(1) the meaning of the idiom is not the result of the compositional function of its constituents;

(2) an idiom is a unit that either has a homonymous literal counterpart or at least individual constituents that are literal, though the expression as a whole is not interpreted literally;

(3) idioms are transformationally deficient in one way or another;

(4) idioms constitute set expressions in a given language;

(5) idioms are institutionalized. (Fernando and Flavell 1981, 17)

Idioms can be metaphorical, but they are different from metaphors such as *a raging storm* where the *storm* is equated to a person who is *raging*, or similes where something is explicitly compared to something else, for example, *this task feels like pulling teeth*. Idioms contain no such explicit comparisons. However, Cruse (1986, 44) states the following: "While idioms and dead metaphors must be distinguished, it should be recognised that they have certain characteristics in common" and he adds that a lot of idioms "began their lives as metaphors". A dead metaphor is a metaphor that has lost the original imagery of its meaning so it is easy to see how a dead metaphor could turn into an idiom. Mäntylä (2004, 36) states that distinguishing dead metaphors from idioms may not be necessary or even possible. Proverbs differ from idioms in that they contain some sort of teaching or "wisdom" whereas idioms do not. Idioms allow people to describe events and situations more vividly with fewer words, saving time and linguistic effort, or they can be used as euphemisms when talking about difficult subject matters (Mäntylä 2006, 153). Because they are conventional, they do not require complex creative strategies like, for example, poetic language does. Idioms have also

traditionally been seen as frozen in the sense that they have a specific structure and specific words that cannot be substituted by other words without compromising the idiomaticity of the expression (Mäntylä 2004, 28). However, some variations are possible.

Mäntylä (2004, 31) presents analyzability and non-compositionality as "two sides of the same coin" as she puts it. "Non-compositionality has often been interpreted as a sign of idioms being dead, that is, their meanings being arbitrary rather than figurative" (ibid.). Even though the meaning is not predictable from the meanings of the individual words, the meaning is not arbitrary: "In fact, as studies on idioms' metaphoricity prove, it is often motivated, that is, a relationship between the metaphorical and literal meanings does exist" (ibid.). Stock et al (1993, 231) argue that "there is a whole class of idioms for which noncompositionality is false". This means that there are idioms that are compositional, and the meaning can be derived from their constituents.

Mäntylä (2004, 32) states that some opaque idioms can be, for example, passivized (*bring home* and *pull strings*, for example). The possibility to reduce and refer through one component to the whole expression is also evident with other fixed phrases (ibid.) The possibility of reduction speaks against frozenness. Analyzability is linked with metaphoricity. Analyzability might seem like a straightforward thing, but it depends on the intuitions of the language user (Gibbs 1993, 63). Gibbs (ibid., 63-64) argues that analyzability is of high importance for understanding and learning idiomatic expressions: if the speaker can comprehend the reasoning for the expression, it will be much simpler to understand and possibly also to remember.

Frozenness refers to invariance in appearance. It has generally been thought of as a characteristic of an idiom, but as stated earlier, idioms can vary. Thus, scholars disagree on the degree of variance in idioms. Moon's (1998) corpus-based idiom study composed of mostly journalistic texts found that variations were surprisingly common. Moon (1998, 123) states that not all variations are caused by the genre and that "very large numbers of FEIs [fixed expressions, incl. idioms] do not have fixed forms, and it would be wrong to claim that they do. The evidence is simply against it." However, transformations can change the meaning or at the very least the nuances attached to the idiom and depending on the context an expression might be understood idiomatically or literally (Mäntylä 2004, 33), regardless of what the intention was. For example, it is literally possible *to shoot oneself in the foot* or it can be an idiom that refers to doing something that later turns out to be harmful to oneself.

Variation allows for creativity in the use of idioms: playing with the relationship between the figurative and literal is possible, and both lexical and syntactic variability add to these possibilities (Mäntylä 2004, 34). It allows for alteration in form or wording to better fit the context and situation while still retaining the idiomaticity of the expression (ibid.) However, idiom transformations can be confusing for a foreign language learner. Two different idioms can get confused with each other or an alteration in the choice of words changes the figurative meaning – often when the phrase is used carelessly – and the mental image created by the idiom becomes comical (Punttila 1994).

Various studies (Fernando 1996, Moon 1998) have shown that idioms are able to undergo some variations. Idioms can, for example, be lexically substituted (*crack the ice* instead of *break the ice*) and syntactically transformed (*breaking the ice*). *Kick the bucket* is often thought of as a totally fixed idiom, but forms *kick the pail* and *kick the can* are possible according to Moon as she has encountered these forms of the expression in real texts with the meaning "to die" (Moon 1998, 123). However, opposing arguments exist. Glucksberg (1993, 21) argues that substituting *crack* for *break* in *crack the ice* is possible, but *break* cannot be substituted with just any verb. *Crush* or *grind* would not be function properly even though they could be regarded as appropriate considering the specific context of speaking about ice. Idiom variations occur in everyday speech but in the media too, as is visible in the examples given by Moon (1992):

> No skeletons in Matthew Taylor's cupboard, they all say. Well, a small door did open behind the Liberal candidate for Truro yesterday and a minor pile of bones was heard to clatter out. (The Guardian, 24 Feb 1987:36)

> A television news reporter asking President Bush: "Did this summit bury the hatchet?"

> Bush: "There is no hatchet" (Moon 1992, 15-16)

Language users do not always agree on the interpretation of an idiom (Mäntylä 2006, 154). Another feature causing confusion, as discussed earlier, is the fact that an idiom can have multiple slightly different forms. For example, in Finnish language, one can either use the expression *mennä kaaliin* or *mennä jakeluun* to signify the same idea: that something has been understood (ibid.). If one is used to using the first form, noticing that the other form is used to signify the same idea is not always without trouble (ibid.). A single idiom can also have multiple interpretations: *joutua ojasta allikkoon* traditionally means ending up from trouble to even worse trouble or someone can interpret it as meaning that the troublesome

situation is made better (ibid., 155). Younger language learners are more accustomed to L2 idioms and can even translate them directly into their L1, causing confusion in other, possibly older or more advanced speakers of the L1.

Loaning idioms from the L2 can cause confusion for the learner himself: *the last straw* is an idiom found in both English and Finnish, but the interpretations differ between the languages (Mäntylä 2006, 156). In English, *the last straw* refers to the biblical quote *about the last straw that broke the camel's back*, and it means the last in a series of unpleasant events that finally makes you feel that you cannot continue to accept a bad situation. In Finnish, however, *the last straw* (viimeinen oljenkorsi) refers to the final possibility of ending a difficult situation. Idioms are ever-changing just like other words: new ones emerge, old ones change or fall out of use completely. The idiom *tehdä konoset* was frequently used in Finnish sports-related news articles and broadcasts in the 1990s but it has since fallen out of use and not many people understand the meaning of it anymore. The idiom's use began as a Finnish athlete called Valentin Kononen became nauseous during a competitive walking race and threw up (Yle 2017). In this case, the idiom was derived from the surname of the person. Also note that here it is evident how idioms are not only used in informal contexts: plenty of them are used in sports-related news articles, for example.

It has generally been presumed that idioms were originally metaphorical, but this view has been challenged. Gibbs (1993, 57) states that "the dead metaphor view of idiomaticity is dead wrong". Gibbs (ibid., 58) states that many scholars have assumed idioms to be dead metaphors since their usage might determine their meaning. For instance, *break a leg* stems from an old superstition in the theater world that wishing a performer good luck would bring him bad luck (ibid.) So, people started wishing bad luck on performers to avoid actual bad luck, thus telling them to *break a leg* (ibid.). Contemporary speakers may comprehend what this idiom means because it is a convention (especially in theater) and not because they know why this idiom means what it does. Furthermore, Gibbs (ibid., 74) states that "the dead metaphor and conceptual views of idiomaticity should not be seen as competing theories." Many idiomatic phrases might actually be dead, or their meanings might be determined arbitrarily due to convention (ibid.)

Even though an idiom's true meaning differs from the literal meaning, the individual words in the idiom can work as a clue to direct the learner into the right interpretation even though this method is not without problems (Mäntylä 2006, 158). For example, the idiom *cost an arm and*

a leg tells the learner that something costs money, or more specifically, costs a lot of money. The word *cost* directs the learner to this interpretation with its literal meaning. Idioms can also create vivid mental images that can guide the learner to the right interpretation. Therefore, idiom dictionaries often contain illustrations of the idioms (Mäntylä 2006, 157-158). Also, the context - words and sentences surrounding the idiom - can work as a tool for understanding the meaning of the idiom. Moon (1998) states the following about context:

> Just as polysemous or homographic words are ambiguous when divorced from the context, so are strings such as *bite the bullet* or *spill the beans*. In context, however, any ambiguity is resolved, with literal interpretations precluded. (Moon 1998, 178)

Context might not always help in comprehension as different people interpret things differently. For example, natives might interpret an idiom differently from non-native speakers. Idioms also vary a lot in their degree of metaphoricity. An idiom such as *kick the bucket*, has a figurative meaning totally different from the literal meaning whereas in an idiom like *to rain cats and dogs* the figurative and literal meaning are not that different as they both refer to actual rain. Etymology can serve as a way to teach idioms. As an example, the idiom *spill the beans* has been claimed to come from an ancient method of voting by using beans. Etymology gives the idiom some context which helps with remembering the idiom and the stories behind an idiom's origin can sometimes be fascinating.

Idioms have generally been thought of as being used only in informal, spoken language rather than in written or formal language (Mäntylä 2004, 34). However, Moon (1998, 69-71) found that idioms are used a lot in newspapers, especially in the headlines. Mäntylä (2004, 34) argues that idioms arouse interest and journalists are thus eager to use them. Idioms are also quite common in Finnish news and sports broadcasts. In his 1982 study "Idioms in English: A Pragmatic Analysis" Jurg Strässler analyzed therapy sessions, trials and White house discussions, and found that regardless of the rather formal settings idioms were used in these types of contexts although less than initially thought (Strässler in Mäntylä 2004, 34-35). The White House conversations especially contained many idioms, but Strässler points out that this might have been due to one person – who was rather low in hierarchy - using a lot of idiomatic language. A person higher up in the hierarchy might be more hesitant to use such informal language as idioms. Strässler also found that idioms are more likely to be used when talking about a third person or an object, but rarely used when the person talks about himself or the listener. The undertones carried by idioms are related to the level of formality (Mäntylä

2004, 35). A lot of words carry undertones, but idioms are almost never neutral in tone so special attention must be directed at the style of the idiom and the situation where it is used. Stock et al (1993, 231-233) point out that while two idioms can have meanings that are close to each other, they might have certain characteristics that make them separate from one another. Stock et al (ibid., 233) also state that the meanings can sometimes be overly simplified. For example, *kick the bucket* is often interpreted with the meaning "to die", but according to Stock et al (1993, 233) "The idiom seems to require death by (more or less) natural causes, and relative suddenly too." This idiom is also more commonly used when the person who died was not close to the speaker so there is a difference in the context of use.

Fernando (1996, 35) divides idioms into three sub-classes: pure idioms, semi-idioms, and literal idioms. Pure idioms are completely non-literal expressions (ibid., 35-36) such as *smell a rat* which has nothing to do with actual rats or smelling. Semi-idioms have one or more literal words and at least one word with a non-literal meaning, often restricted to only that co-occurrence relation (ibid., 36). For example, in *drop names* the verb *drop* has the meaning "to overuse" only when *names* appears together with it. Some semi-idioms allow for lexical variation, for example, *blue film/joke/gag/story/comedian* (blue meaning obscene) (ibid.). Semi-idioms can also be referred to as semi-literal idioms. Literal idioms meet only the invariance or restricted variation criteria of idioms, and they are semantically less complex compared to the two previously mentioned idiom categorizations. *Tall, dark and handsome* and *waste not, want not* are examples of a literal idiom (ibid.) and one can instantly notice that the meaning is easily understood and is thus opaque. Literal idioms meet the invariance or restricted variation criteria of idioms (ibid.).

As one can see from this chapter, idioms have multiple definitions. For the purposes of this study, the idiom definition provided by Mäntylä (2006) will be used since it is not too exclusive or strict and therefore allows for different types of idioms to be discovered in the material collected for the study. A too strict set of criteria might lead to the number and style of idioms discovered being too restricted for any true analysis concerning their different classificiations to be made.

I have now presented what is generally meant by idiomaticity and idioms, and the multiple definitions and classifications of idioms. The use and contexts of idioms were also discussed. I have now also given the idiom definition used for the purposes of this study. Next, I will discuss the role of idioms in second language acquisition.

2.4   Idioms in SLA

Idiom and other formulaic language knowledge is essential to using and understanding a foreign language as "speaking a language with any degree of fluency requires a knowledge of idioms, proverbs, slang, fixed expressions, and other speech formulas. People rarely talk using literal language exclusively" (Gibbs 2010). Idioms can be used to, for example, describe simple matters in a more vivid manner or they can function as euphemisms for difficult subject matters (Mäntylä 2006, 153) The acquisition and comprehension of formulaic language in L2 has not been investigated extensively, even though it is an important part of vocabulary. This might be due to how little attention has been given to vocabulary in general in the field of linguistics (Mäntylä 2004, 81). Most studies on the topic have focused on native speakers, young learners and people who are hard of hearing (ibid., 78). Idioms cause difficulties for L2 learners because they are restricted expressions where the figurative meanings cannot be derived from the literal meanings of the words contained within the idiom. Laufer (1989, 12) discovered in her study of first-year university students based on an error corpus collected over several years that idioms "were translated literally, word by word." Schmitt and Carter (2004, 13) state that even though it is not a simple task to examine the course of formulaic language development in L2 learners, it seems to develop slower than other linguistic aspects. It does not help that idioms are often avoided in speech directed to L2 learners in an attempt to keep speech more understandable for them (Irujo 1986, 236). In addition, L2 learners often come in contact with idioms in non-interactive situations like watching TV or reading, where there is no possibility of negotiating for meaning. The language learner's age plays a part in idiom learning as according to Piaget (1932), children are capable of abstract thinking at around 11-12 years of age. Prior to this age, trying to teach idioms might be a waste of time or at the least, very difficult as the learner cannot possibly comprehend how one can say a thing and mean something completely different.

The significance of L1 in foreign language learning has been studied before (Ringbom 2007 & Jajić Novogradec 2021). Language learning is based on previously acquired knowledge; learners attempt to connect the newly learned elements to the old knowledge. Laufer and Shmueli (1997, 89) state that the results of their study "underscore the importance of attending to newly learnt vocabulary and relating it to the first language." Cross-linguistic knowledge is important when learning new languages, but it "depends on the relationships that can be established between the TL and L1" (Ringbom 2007, 1). If the L1 and the target language (TL) are closely related, pre-existing linguistic knowledge will be of use but not so

much when the languages are distant. Still, Ringbom (2007, 8) states that "grammatical similarities occur even across wholly unrelated languages". Some English idioms do have direct equivalents in Finnish language, for example. Levorato (1993) came to the following conclusion:

> the development of children's ability to produce and understand idioms depends on the development of the same linguistic abilities on which figurative language as well as language in general are based. (Levorato 1993, 101)

The common view is that idioms are difficult for L2 learners. Cooper (1999, 233) argues that it may not even be possible for L2 learners to perfect their knowledge and use of idioms. Arnaud & Savignon (1997, 167-168) support the idea that even advanced L2 learners are not able to achieve native-like competence in using idioms and other "complex lexical units". However, idiom use increases fluency and contributes to native-like competence, and advanced learners can be expected to be able to recognize, interpret and produce some L2 idioms. As discussed in the section on defining idioms, idioms have various characteristics that can affect their comprehension. Idioms are often seen as fully figurative and non-compositional, but the individual words and the imagery created by them can work as a guide on interpreting the meaning. L2 learners need to consider the context in which idioms are present along with the literal meaning of the expression. L2 learners can use various strategies to comprehend the meaning of an idiom. L2 learners may, for example, attempt to guess the meaning from context, use the literal meaning of the components of the idiom as a key to its figurative meaning or refer to an L1 idioms to understand the L2 idiom (Cooper 1999, 243). Nippold and Martin (1989, 59-66) did an experiment where students at different grade levels were presented idioms without any context or as a part of a story. The conclusion was that idioms connected to a story were understood better. However, the context might not always guide to the right interpretation and the same applies to the literal meaning. It can also happen that L2 learners do not recognize the literal meaning and attempting to guess the origins of an idiom could also lead to the wrong interpretation (Mäntylä 2004, 79). Sometimes the relationship between the literal and figurative meanings of an idiom is not clear and close enough, and that is a "stumbling block" for comprehension (Cooper 1999, 244). A Finnish learner of English can easily misinterpret the idiom *blue-eyed boy* as in Finnish the idiom *blue-eyed* refers to a gullible person and in English *blue-eyed boy* refers to a person who is in the position of being a personal favorite of another person.

Irujo (1986) did an experiment with Spanish first language advanced English students. She wanted to study the effect of L1 idiom knowledge in learning L2 idioms. Irujo created three different lists of English idioms: ones that were identical, similar and totally different from the Spanish counterparts. The participants were tested with different types of recognition and production tasks. The results showed that identical idioms were the easiest to recognize and produce. Similar idioms were almost as well recognized and produced but there was an interference effect between the L1 and L2. The completely unsimilar idioms were the most difficult to recognize and produce, but the evidence on negative or positive interference of Spanish idioms was scarce. There was a lot of variance between the test scores, but Irujo points out that the most recognized and correctly produced idioms were those that were used a lot in everyday speech, and which were "opaque" and had uncomplex vocabulary and grammatical form.

Mäntylä (2004) discovered in her study that even advanced Finnish L1 learners of English did not always recognize or understand English idioms. The study was done as a three-stage survey among 144 Finnish L1 English language university students and 36 British English L1 university students or recently graduated. In the survey, the participants were presented idioms with no context as the goal was to find out all the different possible interpretations of the idioms. In the first stage of the study, the participants were asked to rate idiom interpretations given by dictionaries. In the second stage the focus was on finding out the differences in the usage contexts of idioms. In the third and the largest stage, the participants were tasked with giving the interpretation of a certain idiom. It turned out that the Finnish university students had quite a poor knowledge of the meanings of English idioms regardless of them being at an advanced level in their English studies. Idioms with a direct or partial equivalent in Finnish were the easiest to understand but idioms that were completely dissimilar caused difficulties. This supports the previous results of the study by Irujo (1986). Mäntylä also asked about the strategies used for understanding an idiom and it was found that using the L1 for help was the most common strategy. If there was no direct equivalent, the participants tried to focus on one word instead and think of a Finnish idiom that contains the same word. This strategy led to a high number of false interpretations.

3 Material and methods

In this section, I will first present the detailed research questions of the study, and then I will discuss the data and finally, I will discuss the analysis process of the data, going through the steps involved in the process.

3.1 Research questions

The purpose of this thesis was to answer the following research questions:

1. What kinds of differences can be observed between the groups? For example, does one group use more of a certain type of idiom than the other?

2. Is there a link between the reported number of months spent in an English-speaking environment and the number of idioms used in the essays?

The first question focuses on the differences between the two groups: the group with 6 months of experience or less, and the group with more than 6 months of experience. In the second question, the focus is on differences between the participants regardless of the group: the number of months of experience and the number of idioms they have used in their texts. The approach to these research questions is quantitative in an exploratory manner: the limited data does not allow for true quantitative conclusions to be made so instead the focus is on examining quantitative tendencies. The objective is to find out tendencies instead of mathematical correlation since the limited material does not allow for actual statistical analysis. The hypothesis is that the group with over 6 months of experience would use more idioms and that there would be a tendency for higher number of months of experience being linked with more idiom usage, due to Baroto (2017, 2) stating how important exposure to the L2 is for overall English skills.

3.2 Data

The material of this study was collected from The Uppsala Student English Corpus (USE). It is an electronic corpus containing English language essays from the Department of English at Uppsala University. All essays were written by the students during years 1999-2001 (Uppsala Universitetet 2022). The corpus was collected with the aim to create a powerful tool for research into the process and results of foreign language teaching and acquisition. It is available for free online, and it contains a total of 1489 essays written by 440 Swedish

university students of English at three different levels. The average essay is 820 words long, but the first-term essays are a little bit shorter, 777 words on average. The essays cover set topics of different types, some restricted by academic conventions and some more free in style. The essays were written outside of class with a deadline of 2-3 weeks. All of them were written without any supervision or time limits apart from the deadline. The students had unrestricted access to various external source materials such as dictionaries and all kinds of material available on the Internet during the writing process. The students sent the electronic copies of the essays on disk or through e-mail. All 440 students also filled in a questionnaire, answering questions about themselves, including first language, sex and latest English grade among other variables. This provides the opportunity to check the writer's background information for the purposes of this study. All the essays in the corpus are anonymous and a code number is used to refer to different students instead of names. This also allows for finding the different essays produced by the same student if required.

Different corpuses were considered and compared, but none of them had the features necessary for carrying out the study. First, the corpus had to be a written learner corpus with English as the L2 and the style of the texts had to be non-academic to allow for a freer writing style that makes using idioms possible. And second, the writers had to be at a high enough level in their language skills so that they have the ability to use idioms in their texts. A spoken corpus was also considered as spoken language often contains more idioms, but this would also depend on the types of tasks used for the corpus material and how comfortable the students are when they discuss having the knowledge that they are being recorded, so a written corpus was deemed a better fit for the study.

The Uppsala Student Corpus was downloaded from the Internet along with the instructions on how to use it and a database in Microsoft Excel containing the background information on all the participants. Fifty texts from the corpus were chosen for inspection. Each of them was chosen from the culture course assignment which was the last assignment of the first term. The topics included, for example, slavery in the USA and women's rights to education in the early 20th century. It could be argued that topics related to culture can evoke more usage of idioms, thus making these types of essays a good choice for this study. It also helps that the essay requirements did not call for the use of academic writing style which allows for more informal language and thus, an opportunity to use idioms freely. Half of the texts (25 texts) were from those participants who had reported zero to six months of experience in an English-speaking environment and the other half from those who had reported over 6 months of

experience. This division allows for comparing the idiom content of the texts of those with a lot of experience and those with little or no experience. The amount of experience varied drastically, from 0 months to as much as 204 months.

Majority of the students chosen for this study had Swedish as their first language and most of them were female in both groups. Only 2 out of the 50 students had a first language other than Swedish and only 12 were male. One of the Swedish language students also had Finnish as their first language, though. This is explained by the university being in Sweden and the fact that most university students are female. The average time the essay writers had spent in an English-speaking country was 1,78 months in the 0 to 6 months group and 27,08 months in the over 6 months group. This is as expected as the over 6 months group was not given any maximum limit for the time of experience. As mentioned earlier, the students in the over 6 months group could have as much as 204 months of experience. The students with such a high number of months of experience had often lived in an English-speaking country for a long period. Some had acquired a high amount of experience by having an English-speaking partner or by travelling a lot. Exchange students were also common.

The number of idioms found in the essays was lower than expected when taking into accord the fact that they were English students who can be expected to have good English skills and many years of English studies behind them, and thus a good grasp on idiomatic language and the willingness to use it. The criteria for idioms were kept rather lenient for the purposes of this study but the discovered number of idioms remained low regardless of that fact. The language used in the essays could be described as simple and basic. The writers were probably trying to avoid taking unnecessary risks and this also includes avoiding idiomatic language as there is a risk of using it in an inappropriate manner or context. It is unfortunate that the discovered number of idioms was so low as it does not truly allow for drawing any true quantitative conclusions as to whether the participants with more experience in an English-speaking country use more idiomatic language compared to those participants with only little experience. However, the present study is exploratory in nature, so an extensive amount of data is not a requirement.

3.3 Analyzing the data

Due to the drastically varying nature of idioms, using any type of search function to look for them in the essays was literally impossible. The only option was to read the essays manually in a very careful manner, checking for idiomatic expressions, counting them, and marking the

exact written form down in a table, along with the type of the idiom and the rest of the information related to the text such as the participant ID and the number of months of experience. This method was tedious and time-consuming. It must also be stated here that this method is prone to error as the reading was done by a human being and not by a computerized electronic system. However, as stated earlier, the nature of idioms effectively prevents the use of a computer system, and the essays were read multiple times carefully to minimize the risk of human error. The three types of idioms used for the idiom type classifications were literal, semi-literal and pure, based on Fernando (1996). This type of classification fits the purposes of this study as the lenient criteria for the idioms allows for all those subclassifications of idioms to be found in the texts. Using stricter criteria for idioms might not necessarily allow for discovering all of the three subclassifications.

The criteria of idioms for the purposes of this study were not too strict and followed the ideas of Mäntylä (2006). An idiom, for the purposes of this study, is defined in the following manner:

- figurative expression

- two or more words

- the true meaning of the expression differs from the literal meaning

- is conventional.

The last criterion, conventionality, can be argued as being essential to this study as any kind of so called "insider idioms" which are used and understood in restricted groups of people, will not be recognized as idioms at all or at least the meaning will not be understood. One could even make the case that an idiom is an idiom only when it is understood and used by the general public. This is in accordance with Häkkinen (2002, 451) who states that idioms are conventional. However, it must be noted that conventionality can at times be controversial and an idiom might not be found in a dictionary even though a person can easily recognize and understand the idiom when he sees it. The sub-class of the idiom is classified according to the classification by Fernando (1996, 35): the idiom is assigned into the literal, semi-literal or pure idiom group.

The total number of idioms, the number of each idiom type (literal, semi-literal or pure) and the exact written forms of the idioms for each essay were marked down in the table. One

essay had to be read and checked multiple times to minimize the risk of missing an idiom by accident. Idioms found in quotes were not counted since quotes are not truly produced by the essay writer himself. There is a possibility that the writers used idioms they had read in the texts they had read in preparation for the essay in their own writing, but one can argue that this is not a problem since using the idiom correctly in the appropriate context requires more skill apart from just copying the idiom. The writer needs to possess the knowledge of when and in what context to use the idiom. Since formulaic language is learnt as a single unit despite the unit being a string of multiple words, it can be argued that learning an idiom by reading another person's text and then using the idiom in your own text is no different from learning a new single word in the same manner. Using a single word correctly also requires more skill than just "copycatting" it. Furthermore, small mistakes in the exact form of the idiom were not seen as a basis for disqualifying it as an idiom if the idiom remained understandable. For example, if the writer wrote *beat around a bush* instead of *beat around the bush*, the idiom would be counted as substituting *the* for *a* does not drastically diminish the understandability of the expression even though one could argue the idiom is incorrectly written as it differs from the common form. However, as stated earlier, research shows that idioms can be altered in some ways even though fixedness is often seen as a basic component.

4   Analysis

In this section, I will first go through the overall results of the study. Then I will move on to analyzing the idioms found in the essays. The idioms will be presented in groups according to their classification (literal, semi-literal and pure) under different subchapters. I will also present the context of the idioms with quotes and discuss the interpretations. Furthermore, I will be commenting on possible alterations in the idioms and discussing how the idiom still fulfills its function even if it differs from the common form. The common forms and interpretations of idioms have been searched from well-known and reputable dictionaries such as the Cambridge Dictionary.

A total number of 41 idioms were discovered in the fifty essays. The average number of idioms per essay was thus 0,82 idioms. It must be noted, however, that 26 of the essays, over half of the total number, did not contain any idioms whatsoever. This means that the 41 idioms were found in the 24 essays that did contain idioms. 11 of the essays in the 0-to-6-month group contained idioms and 13 of the essays in the over 6 months group contained idioms. 16 idioms were found in the essays written by those participants with 0 to 6 months of experience in an English-speaking environment and 25 in the essays written by participants with over 6 months of experience. This means that 61% of the total idioms discovered were in the group with over 6 months of experience and 39% were in the group with 0 to 6 months of experience. Average number of idioms per essay was 0,64 in the 0 to 6 months group and 1,0 in the over 6 months group. The 0-to-6 months group contained 1 literal idiom, 7 semi-literal idioms and 8 pure idioms. The over 6 months group contained 4 literal idioms, 8 semi-literal idioms and 12 pure idioms.

4.1   The literal idioms in the essays

Compete on the same terms was the only literal idiom used by the students in the 0-to-6 months group and it was discovered in an essay concerning slavery in the US: "Lastly, by being inferior to the white race, he could not compete on the same terms with the white man". According to Collins Dictionary (2022, s.v. compete on the same terms) "If two people or groups compete on equal terms or on the same terms, neither of them has an advantage over the other." The idiom shows that the life possibilities between black and white men were unequal. It is classified as a literal idiom as the meaning is literal. Doing well in life is commonly seen as a competition, not explicitly, but a lot of people compare themselves to

others in terms of education and wealth, for example, and these things are seen as something that divides people into different groups.

The over 6 months group used 4 literal idioms in their essays, and I will now present them one by one. The first one was *to tackle*, discovered in a text concerning Robert Owen:

> When political issues arose in New Harmony, on account of different opinions regarding the government of the community, Owen didn't have the knowledge to tackle them and remain in control of his project.

To tackle refers to solving a problem. The idiom is used to convey the idea that Owen was not able to solve the political issues due to his lack of knowledge. The idiom is classified as literal since the meaning is quite opaque especially considering the context. The second literal idiom was *let loose*: "They do this by stating that if a black man was let loose and exposed to free competition, he would be outwitted and become a burden to society." *Let loose* means to allow someone or something to move or go freely, without constraints. It is classified as a literal idiom since the meaning conveyed by *loose* makes the meaning opaque. The expression conveys an image of a person not chained down or constrained by anything. Note that the quote also contains the semi-literal idiom *a burden to society* but that will be discussed in the next chapter.

The third literal idiom in the over 6 months group's essays was *to put an end to*: "The authorities tried to put an end to sweatshops for a long time". The idiom is clearly literal as the correct meaning is easily interpreted. The fourth and final one was *not without a fight*: "…basically the whole continent were ruled by white people and the Indians was put on reservations (1851). But not without a fight, the Indians fought hard to keep their land". The idiom is classified as literal since the meaning is opaque and literal as the Native Americans did fight.

4.2   The semi-literal idioms in the essays

Now I will present the semi-literal idioms found in the essays, starting with the idioms in the 0-to-6 months of experience group. The group used 7 semi-literal idioms. *In the long run* was the first semi-literal idiom discovered in the group's essays. The essay dealt with women's education in the 1800s and the exact quote goes "But the increased amount of highly educated women would in the long run prove that the opponents were wrong in all aspects." According to Cambridge Dictionary, the idiom refers to "a time that is far away in the future". This idea

could have been stated in a non-idiomatic manner too while remaining rather compact, but the use of an idiom makes the text more vivid and interesting to read.

Burning issue was the second semi-literal idiom found in the group's essays. "In the modern society of today women's education is not a burning issue, it goes without saying that education is a good thing for both men and women." Collins Dictionary (2022, s.v. burning issue) says the following about the meaning of this idiom: "A burning issue or question is a very important or urgent one that people feel very strongly about." The writer wants to state in a compact and fluid manner that the issue of women's education is no longer an issue that is debatable, even though it once used to be.

Push the boundaries was found in an essay about the treatment of Native Americans: "They did sign treaties peacefully to start with, but were constantly disappointed by their new white neighbours who kept pushing the boundaries". According to English Logica (2022, s.v. push the boundaries), "If you push or expand beyond the boundaries, frontiers or limits of something or somewhere you aim to extend beyond the established norms or constrains of something." It could be argued that the idiom generally refers to pushing boundaries in a good sense as in coming up a new technological innovation for example, but it probably can stand for negative boundary-pushing too, so the idiom here was counted.

Two semi-literal idioms were discovered within the same essay: *a poor school for* and *darken someone's mind*. "Mr Douglass said that slavery was a poor school for the human intellect and heart and that there was no natural inferiority between blacks and whites". *A poor school* refers here to the detrimental effects that slavery causes in humans. "Many attempts had been made to reduce their moral and darken their minds but they had all failed". *Darken their minds* refers to reducing someone's moral and lowering their willingness to stand up for themselves.

No high thoughts and *take a fall for the worse* were discovered in an essay about colonialism. "There were no high thoughts of the Indians" and" The Indians' reputation also took a fall for the worse when several Indian attacks on colonists were made" are the exact quotes from the essay. *No high thoughts* refers to having a low opinion, instead of a high opinion, of something. *Take a fall for the worse* is an expression used when a situation gets worse. Both idioms are rather easy to interpret correctly yet they contain a word that is non-literal, so they are classified as semi-literal idioms.

Now, I have presented all the idioms found in the essays by the 0-to-6 months group and next I will present the semi-literal idioms found in the essays by the group with over 6 months of experience. There were 8 semi-literal idioms in the group's essays. *A burden to society* was mentioned in the previous chapter about literal idioms and it was used in an essay about slavery. This is a semi-literal idiom that refers to something being troublesome to something, in this case to society. It is classified as semi-literal because the word *burden* is non-literal, but *society* is a literal word in the idiom.

Competition between man and machine was an idiom used in an essay about Robert Owen: "As the competition between man and machine hardened, with great unemployment as an effect, Owen's attention went toward cooperatism." The expression is a semi-literal idiom as the *competition* is a non-literal constituent whereas *man and machine* refers to actual humans and machines. The expression refers to new machines in factories making the human workforce redundant and leading to unemployment; there is sort of a power struggle between the human workforce and the machines.

Have a negative ring to was a semi-literal idiom discovered in an essay about Native Americans and their treatment: "However, even though most of these titles have a definite negative ring to them, they were often followed by a term of respect of some sort." The idiom refers to a word or expression having a negative overtone, "ring". Since the word *ring* is to be understood non-literally here, the idiom is classified as semi-literal. *A matter of course* was used in an essay about women's right to education: "The situation for the right for women to get education is still not a matter of course for the majority of the worlds women." The proper idiom proposed by the Cambridge Dictionary (2022, s.v. as a matter of course) is *as a matter of course* and it refers to "a usual part of the way in which things are done and is not special". The quote means that women's right to education is still not a usual thing in most of the world. Even though the idiom in the essay does not follow the proper form proposed by the Cambridge Dictionary, it remains recognizable and is easily interpreted correctly. It is semi-literal as the word *course* is non-literal while *matter* is literal.

A chilly relation was used in an essay about Robert Owen: "…Owen had no solution to the problems, since he himself wanted control and personally had a chilly relation towards religion." *Chilly relation* refers to having a relation that is less than positive while not quite negative. Collins English Dictionary (2022, s.v. chilly) states that "You say that relations between people are chilly or that a person's response is chilly when they are not friendly,

welcoming, or enthusiastic." So basically, Owen was not in favor of religion and preferred to keep some distance from it according to the essay. The idiom is semi-literal as it consists of the non-literal *chilly* and the literal constituent *relation*.

To be better off was used in two separate instances in an essay about slavery: "It was namely argued that the slaves on the plantations were much better off than the free ones in the North" and "Slaves were now treated well and were better off both morally and physically under supervision by caring masters than they would be as free labourers." According to Cambridge Dictionary (2022, s.v. better off), *better off* can refer to either having more money or being in a better situation. In the essay, it refers to the latter. *To be better off* is semi-literal: the word *better* is literal and has this meaning only when it co-occurs with *off*.

To fit in was used in an essay about women's education: "These people considered education as something important since it made it easier for them to fit in into their new environment." The concept of *fitting in* (Cambridge Dictionary 2022, s.v. fit in) is that you get comfortable in a new group and become accepted as part of it. The expression is a semi-literal idiom as it contains the non-literal constituent *fit* and the literal constituent *in*. It is also restricted in meaning in sense that *fit* only has this meaning when it co-occurs with *in*.

4.3   The pure idioms in the essays

I will first present the pure idioms in the essays by the 0-to-6 months of experience group and then the other group with over 6 months of experience. The 0-to-6 months group used 8 pure idioms in their essays. *An Eden on Earth* was found in an essay about the discovery of America. "The Europeans had for a long time believed that an Eden on earth lay somewhere in the west, and thought they were going to discover just that in America." The essay discussed the American colonists. The idiom used here portrays the idealized thinking the colonists had about the new continent they discovered. The more commonly used form of this idiom is known as *Heaven on Earth* (Merriam-Webster 2022, s.v. heaven on earth) but *Eden* fulfills the same function here, so the substitution does not make it a non-idiom even if it is less used than the commonly known form. The idiom is classified as pure since it is fully non-literal.

An appetite for blood was discovered in an essay about Native Americans. "Indians are said to have a natural instinct to kill and plunder and to have an appetite for blood." The more commonly known idiom is *bloodthirsty* (Cambridge Dictionary 2022, s.v. bloodthirsty) but

appetite for blood creates the same mental associations and thus the idiom functions as an idiom and is recognizable as one even if it is less commonly used than *bloodthirsty*. The idiom is classified as pure.

An essay about discovering America was found to have three different instances of pure idiom usage. *Carve it out of* was the first one: "They wanted to create a world, carve it out of the wilderness and polish it up." This quote also contains another idiom found in the essay, *polish it up*. It must be noted that *create a world* might also seem to be an idiom, but it is not since it is not conventional. It is idiomatic, but it is not a true idiom because it lacks the conventionality. *Carve it out of* might not seem idiomatic at first but one needs to take into account the context: physically carving something out of the wilderness is not something you could do in reality. According to the Cambridge Dictionary (2022, s.v. carve out sth), *carve out something* is the more commonly used idiom, and it refers to creating or obtaining something that helps you by skillful activities. For example, "she carved out a reputation for herself as a high-powered lawyer". In the essay, the meaning is that they would *carve a world out of the wilderness*. *Polish it up* refers to improving something. The third and final idiom in the essay was *strengthen the grip*: "This, however, only strengthened the grip that religion had on people." Collins English Dictionary (2022, s.v. tighten the grip) states that the proper form of this idiom is *tighten the grip* but *strengthen* functions too and the idiom remains recognizable.

Open the way was discovered in an essay where the topic was women's education: "The new demand for qualified teachers opened way for women to enter into higher education." The missing article *the* was not an issue here as the idiom remains recognizable. Just like before, using an idiom here keeps the text more interesting as opposed to using an expression like *made it possible for*. It is classified as a pure idiom since it is totally non-literal.

Carry the scars was an idiom used twice in slightly different forms within the same essay about child labor: "These children suffered badly and those surviving to adulthood often carried the scars of their early exploitation." and the second one "They were all carrying the scars of their early exploitation." The idiom creates a gruesome image of the effects of child labor. Cambridge Dictionary (2022, s.v. bear the scars) recognizes the idiom in the form *bear the scars* but substituting *bear* for *carry* does not change the meaning or recognizability of the idiom. The idiom is interpreted by Cambridge Dictionary (ibid.) as "to still suffer emotional pain from something unpleasant that happened in the past". It remains unclear from context

whether the writer intended to refer to physical or mental effects left by the childhood labor, but it can be safely assumed that the writer could refer to both. The idiom is classified as a pure idiom as it is fully non-literal.

Now I have presented all the idioms in the first group's essay and will move on to the second group. The group used 12 pure idioms in total. The idioms *A very quick climb up the societal ladder* and *drop out* were used in an essay about Robert Owen: "Robert Owen who was born in 1771 made a very quick climb up the societal ladder when he in early years dropped out of school and set of to work in London." *Climbing up the societal ladder* refers to improving one's status in society. The idiom is pure as it is totally non-literal: in reality, nobody is climbing anything and there definitely is no ladder. *Drop out* means to quit something, in this case, school. People who have quit school can also be referred to as *high school dropouts* or so. The idiom is pure as it is totally non-literal. *Make ends meet* was used in an essay about women's education: "Her husband gave her the money he earned, and it was up to her to try to make ends meet." The Cambridge Dictionary (2022, s.v. make ends meet) explains the idiom's meaning as "to have just enough money to pay for the things that you need". The idiom is fully non-literal and therefore pure.

A gift from God was used in an essay about the battles between Native Americans and the settlers: "They were the chosen people, the land was a gift from God." Merriam-Webster (2022, s.v. a gift from the gods) recognizes the form *a gift from the gods*, which refers to "a very lucky and helpful thing". *A gift from God* functions properly here, too. The idiom is clearly pure. The pure idiom *the Angel in the House* was discovered in an essay about women's education: "The women were seen as the Angel in the House responsible for the sacred of the family". *Angel in the House* refers to the woman as a housewife who is responsible for the upkeep of the family home. The idiom's origins are in a poem written by Coventry Patmore in 1854. The idiom is classified as pure since it is non-literal. The same idiom was used in another essay about the same topic but in the slightly different form *the Angels of the house*:

> The reforms naturally effected the situations of "the Angels of the house"
> (referred to the women whose duty were the household and the children), the
> women's position and education slowly grew better.

Making the expression plural and substituting *in* for *of* does not change the understandability or meaning of the idiom in this case. The idiom also remains pure as it is completely non-literal.

Turn out was used three times in the same essay. The first two of them were used in the same sentence:

> Owen's experiments in New Lanark and New Harmony turned out very differently, and in this essay the two experiments will be compared with each other, considering similarities and differences concerning how they both started, how they developed and how the results turned out.

The third occasion of *turn out* was the following: "in New Harmony he had started with a community and tried to make it productive, which had turned out to be a more difficult aim than he probably had anticipated." *Turn out*, according to the Cambridge Dictionary, has the meaning "to happen in a particular way or to have a particular result, especially an unexpected one". The idiom is pure as it is totally non-literal.

To put up with was used in an essay about slavery in the US: "While being a slave one had to put up with being treated as a piece of property, and it was up to the master to do whatever he wanted to do to his slaves." According to Cambridge Dictionary (2022, s.v. put up with sth/sb), *put up with sth/sb* means "to accept or continue to accept an unpleasant situation or experience, or someone who behaves unpleasantly". The idiom is pure as it is completely non-literal.

The two final examples of pure idioms in the over 6 months of experience group are from the same essay which had women's right to education as the topic. The first pure idiom is *to be on the safe side*:

> To be on the safe side and to once and for all prove that the size of the brain was an important factor when determining a persons intelligence quotient craniometry, that is cranial measurement, was used.

According to Cambridge Dictionary (2022, s.v. (just) to be on the safe side), *(just) to be on the safe side* means "being especially careful in order to avoid something unpleasant". In this context, "something unpleasant" would be the situation where people were to find out that the size of the brain is not a valid measurement of intelligence. The idiom is pure because it is non-literal in full: there is no actual "safe side". The second pure idiom in the essay was *if the worst came to the worst*: "Studying could cause illness and if the worst came to the worst, it

could also lead to that a woman no longer would be able to have children." According to Cambridge Dictionary (2022, s.v. if the worst comes to the worst), *if the worst comes to the worst* means that "if the situation develops in the most serious or unpleasant way". In this context it means that the most serious consequence of women studying would be infertility. The idiom is a pure idiom as any part of it cannot be interpreted in a literal sense.

4.4 Discussion

Now that I have presented the idioms found in the essays along with their interpretations and the context, I will discuss some of the results of the analysis taking into accord the research questions. First and foremost, it is evident that the group with over 6 months of experience in an English-speaking environment used more idioms and they used more of the different types of idioms, including literal ones. The 0-to-6 months group used 1 literal idiom, 7 semi-literal idioms and 8 pure idioms in their essays. The over 6 months group used 4 literal idioms, 8 semi-literal idioms and 12 pure idioms in their essays. The differences were not that drastic, as one can see. It does show a tendency for the students in the group with over 6 months of experience using more idiomatic language, though. Also, the group with over 6 months of experience used more pure idioms which are more difficult to interpret than literal or semi-literal ones. Because of this, it is surprising that the group with less experience used only one literal idiom. But on the other hand, the other group did not use that much more of them, and it could be that literal idioms are somewhat rare considering that most idioms are metaphorical and at least somewhat non-literal even if not completely non-literal.

Looking at these results, it seems that Marton (Marton in Mäntylä 2004, 21-22) was correct in saying that being exposed to phrasal language such as idioms in L2 is not sufficient to increase even an advanced learners' knowledge of it. However, Vaurio (1998, 47) stated that incidental vocabulary acquisition has a connection with language proficiency in the way that better knowledge of a language leads to more effective incidental vocabulary acquisition and the participants of this study can be assumed to have good language proficiency as they are studying English at a university. However, the data does not reveal how old the participants were when they spent time in an English-speaking environment. The age during the time of the stay in an English-speaking environment very likely has an effect on how well the student is able to learn incidentally. The results of this study are in line with Mäntylä's (2004) observation that students often have trouble with idioms: if they cannot comprehend them, they also cannot use them in their own texts.

When looking at the number of months spent in an English-speaking environment in the over 6 months of experience group, it appears that there is no tendency for higher numbers of months being linked with more idiom usage. For example, the student with 204 reported months in an English-speaking environment used only one idiom in her essay whereas students reporting 7-15 months used 2-4 idioms in their essays. It must be noted that a lot of the essays in the over 6 months group used no idioms at all and the amounts of experience in these instances ranged from 10 to 90 months. There is a similar tendency in the 0-to-6 months group too: a student with 0 months of experience used 2 idioms in her essay and another student with 2 months experience used 3 idioms. The students with 5 or 6 months of experience used 0-2 idioms in their essays. One must remember that a lot of the students did not use any idioms, though, regardless of their amount of experience. It does seem like there is no link between higher usage of idioms and higher number of months of experience in an English-speaking environment. In fact, it seems that the link is opposite with more experience leading to less idiom usage. It seems counter-intuitive as one could easily assume that more exposure would lead to better overall skills in English and therefore more idiomatic language, as stated by Baroto (2017, 2).

There might be some kind of a threshold around 6 months of experience where the experience leads to higher idiom usage, but the effect diminishes at some point as the amount of experience goes higher. Perhaps it is the context of a course essay that makes the students avoid taking risks and thus avoiding using idiomatic language, but it does not explain why the effect is more pronounced with more experienced students. It might be that experienced students have overall better English skills and higher grades, and they avoid taking risks to keep their essay grades high. It must be noted, however, that the material of this study is very limited and making any true quantitative conclusions would require a higher number of essays to analyze.

5   Conclusion

I have now presented some theoretical background for the purposes of the study, discussed the data and methods used in it and analyzed the findings of it. I began the theoretical section by discussing second language vocabulary acquisition and then moved on to formulaic language. Then I discussed the various definitions of idioms and then finally talked about idioms in second language acquisition. In the material and methods section I presented the data and how it was collected, and the methods I used in analyzing the data. In the analysis section I presented all the idioms found in the essays, grouping them into literal, semi-literal and pure idioms.

The purpose of this study was to find out what kinds of differences in idiom usage does the amount of experience in an English-speaking environment lead to, and whether there is a link between how many months the student has spent in an English-speaking environment and the number of idioms the student used in the essay. The answer to the first research question is that there was a difference between the groups, with the group with over 6 months of experience using more idioms of all types, especially the pure idioms which are thought of as more difficult. The answer to the second research question is that higher reported numbers of months spent in an English language environment seemed to be linked with less idiom usage in the essays. It could be that 6 months is a sort of threshold for enhancing idiom usage and the effect diminishes as experience increases beyond this point, even though it does not explain why the students reporting a high amount of experience used even less idioms than a lot of the students reporting under 6 months of experience. This idea needs more research and could be researched with a wider quantitative study in the future as the present study had only limited data. It might also be the case that the students are attempting to avoid taking risks with their language use in the essays, since the essays will be graded, and this includes avoiding idiomatic language. It could be that the students with a lot of experience and good overall English skills (and good grades) have learned to avoid taking risks in essay writing to keep their essay grades high, leading to texts that are grammatically correct but lack idiomatic language that would make the text more interesting.

Considering how important idioms are for achieving a high, native-like language fluency, the topic of idiom learning in a second language should be given more attention in SLA research. Even if incidental learning and exposure to the second language turn out to be ineffective for learning idioms, researchers and teachers can always consider the various other methods

available for teaching idiomatic language. What is certain is that idioms make a language more vivid and interesting, and therefore they should not be forgotten about. Idioms also keep language compact, expressing ideas that would otherwise require multiple sentences in just a few words. Thus, idiom knowledge is crucial for good writing and speaking skills, and idioms should be given more focus in teaching, especially at higher levels of education, but also in earlier stages to sort of "prime" the students for acknowledging idiomatic language in the texts they read.

References

Primary source

University of Oxford. 2003. *The Uppsala Student English Corpus (USE)*. Oxford Text Archive. Accessed 1 September 2021. http://hdl.handle.net/20.500.12024/2457

Secondary sources

Arnaud, P. J. L., & Savignon, S. J. 1997. "Rare words, complex lexical units and the advanced learner." In: *Second language vocabulary acquisition*, edited by James Coady & Thomas Huckin, 157-173. Cambridge: Cambridge University Press.

Baroto, Muhammad Ali Agung. 2017. "The Effects of Language Input, Learning Environment, and Motivation Toward Second Language Acquisition." *Linguistics, literature and english teaching journal* 6(2). Accessed 26 April 2022. https://jurnal.uin-antasari.ac.id/index.php/let/article/view/1456

Brinton, Donna M., Marguerite Ann Snow, and Marjorie Bingham Wesche. 1989. *Content-Based Second Language Instruction*. New York: Newbury House Publishers.

Cambridge Dictionary. 2022. "Cambridge Dictionary." Accessed 28 March 2022. https://dictionary.cambridge.org

Coady, James. 1997. "L2 vocabulary acquisition through extensive reading." In: *Second Language Vocabulary Acquisition*, edited by James Coady and Thomas Huckin, 225-237. Cambridge: Cambridge University Press.

Collins Dictionary. 2022. "Collins English Dictionary." Accessed 28 March 2022. https://www.collinsdictionary.com/dictionary/english

Cooper, Thomas C. 1999. "Processing of Idioms by L2 Learners of English." *TESOL Quarterly* 33, issue 2: 233-262.

Council of Europe. 2022. "Qualitative aspects of spoken language use - Table 3 (CEFR 3.3): Common Reference levels." Accessed 30 March 2022. https://www.coe.int/en/web/common-european-framework-reference-languages/table-3-cefr-3.3-common-reference-levels-qualitative-aspects-of-spoken-language-use

Cowie, A. P. 1998. "Introduction". In: *Phraseology: Theory, Analysis, and Applications*, edited by A. P. Cowie, 1-20. Oxford: Clarendon.

Cruse, David Alan. 1986. *Lexical Semantics*. Cambridge: Cambridge University Press.

Ellis, Rod. 1994. *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

English Logica. 2022. "English Logica." Accessed 28 March 2022. https://englishlogica.com

Fernando, Chitra. 1996. *Idioms and idiomaticity*. Oxford: Oxford University Press.

Fernando, Chitra, and Roger Flavell. 1981. *On Idiom: Critical Views and Perspectives*. Exeter: University of Exeter.

Gibbs, Raymond W. Jr. 1993. "Why idioms are not dead metaphors?" In: *Idioms: processing, structure, and interpretation*, edited by Cristina Cacciari and Patrizia Tabossi, 57-77. Hillsdale: Erlbaum.

———. 2010. "Idioms and Formulaic Language." In: *Oxford Handbook of Cognitive Linguistics*, edited by Dirk Geeraerts and Herbert Cuyckens. New York: Oxford University Press. Web. Accessed 13 of April 2022 https://www-oxfordhandbooks-com.ezproxy.utu.fi/view/10.1093/oxfordhb/9780199738632.001.0001/oxfordhb-9780199738632-e-27

Glucksberg, Sam. 1993. "Idiom Meanings and Allusional Content." In: *Idioms: processing, structure, and interpretation*, edited by Cristina Cacciari and Patrizia Tabossi, 3-26. Hillsdale: Erlbaum.

Häkkinen, Kaisa. 2002. "Idiomi etsii paikkaansa." *Virittäjä* 106, no. 3: 450-457.

Irujo, Suzanne. 1986. "Don't put your leg in your mouth: Transfer in the acquisition of idioms in a second language." *TESOL Quarterly* 20, 287–304.

Jajić Novogradec, Marina. 2021. "Positive and Negative Lexical Transfer in English Vocabulary Acquisition." *ELOPE (Tiskana izd.)* 18, no. 2, 139-165.

Järvinen, Heini-Marja. 2014. "Kielen opettamisen menetelmiä" [The Methods of Language Learning]. In: *Kuinka kieltä opitaan* [How language is learned], edited by Päivi Pietilä and Pekka Lintunen, 89-113. Helsinki: Gaudeamus.

Katz, Jerrold J., and Paul M. Postal. 1964. *An Integrated Theory of Linguistic Descriptions*. Cambridge: The MIT Press.

Krantz, Gösta. 1991. *Learning Vocabulary in a Foreign Language: A Study of Reading Strategies*. Göteborg: Acta Universitatis Gothoburgensis.

Laufer, Batia. 1989. "A factor of difficulty in vocabulary learning: Deceptive transparency." *AILA Review* 6, 10-20.

Laufer, Batia and Karen Shmueli. 1997. "Memorizing New Words: Does Teaching Have Anything To Do With It?" *RELC Journal* 28, no. 1: 89-108.

Levorato, Maria Chiara. 1993. "The acquisition of idioms and the development of figurative competence." In: *Idioms: Processing, structure, and interpretation*, edited by Cristina Cacciari and Patrizia Tabossi, 101-128.

Merriam-Webster. 2022. "Merriam-Webster." Accessed 28 March 2022. https://www.merriam-webster.com

Moon, Rosamund. 1992. "Textual aspects of fixed expressions in learner's dictionaries." In: *Vocabulary and applied linguistics*, edited by Pierre J. L. Arnaud and Henri Béjoint, 13-27. London: MacMillan.

———. 1998. *Fixed Expressions and Idioms in English: A Corpus-Based Approach*. Oxford: Clarendon.

Mäntylä, Katja. 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*. Jyväskylä: University of Jyväskylä.

———. 2004. *Introducing Marton*. Jyväskylä: University of Jyväskylä.

———. 2004. *Introducing Strässler*. Jyväskylä: University of Jyväskylä

———. 2004. *Introducing Weinert.* Jyväskylä: University of Jyväskylä.

———. 2006. "Vieraan kielen idiomit – banaaninkuoria vai helmiä?" [Foreign language idioms – banana peels or pearls?] In *Kielen päällä: näkökulmia kieleen ja kielenkäyttöön* [On the tip of the tongue: viewpoints on language and language use],

edited by Riikka Alanen, Hannele Dufva and Katja Mäntylä, 151–164. Jyväskylä: Jyväskylän yliopisto, Soveltavan kielentutkimuksen keskus.

Nagy, William. 1997. "On the role of context in vocabulary learning." In: *Vocabulary: Description, Acquisition, and Pedagogy*, edited by Norbert Schmitt and Michael McCarthy, 64-83. Cambridge: Cambridge University Press.

Nation, I. S. P. 2001. *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Nattinger, James R., and Jeanette S. DeCarrico. 1992. *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.

Nenonen, Marja. 2002. *Idiomit ja leksikko. Lausekeidiomien syntaktisia, semanttisia ja morfologisia piirteitä suomen kielessä* [Idioms and the lexicon. The syntactic, semantic and morphological features of phrasal idioms in Finnish language]. Joensuu: Joensuun yliopisto.

Niitemaa, Marja-Leena. 2014. "Kuinka vieraan kielen sanoja opitaan ja opetetaan" [How foreign language words are learned and taught] In: *Kuinka kieltä opitaan* [How language is learned], edited by Päivi Pietilä and Pekka Lintunen, 138-164. Helsinki: Gaudeamus.

Nippold, Marilyn A. & Stephanie Tarrant Martin. 1989. "Idiom interpretation in isolation vs. context: A developmental study with adolescents." *Journal of speech and Hearing Research* 32, 59–66.

Oxford Learner's Dictionaries. 2022. "Oxford Learner's Dictionaries." Accessed 13 April 2022. https://www.oxfordlearnersdictionaries.com

Piaget, Jean. 1932. *The language and thought of the child*. London: Routledge and Kegan.

Pietilä, Päivi, Katalin Doró and Renata Pípalová. 2015. *Lexical Issues in L2 Writing*. England: Cambridge Scholars Publishing.

Punttila, Matti. 1994. "Voittajan käsi kohoaa salkoon. Kielikuvien salakarikoista" [The winner's hand rises to the pole. About the secret pitfalls of figures of speech]. *Kielikello* [Language bell] 1. Web. https://www.kielikello.fi/-/-voittajan-kasi-kohoaa-salkoon-kielikuvien-salakarikoista

Ringbom, Håkan. 1987. *The role of the first language in foreign language learning*. Clevedon: Multilingual Matters.

———. 2007. *Cross-Linguistic Similarity in Foreign Language Learning*. Clevedon: Multilingual Matters

Schmitt, Norbert and Ronald Carter. 2004. *Formulaic Sequences: Acquisition, Processing, and Use*. Philadelphia: John Benjamins Publishing Company.

Stock, Oliviero, Jon Slack, and Andrew Ortony. 1993. "Building Castles in the Air: Some Computational and Theoretical Issues in Idiom Comprehension." In: *Idioms: processing, structure, and interpretation*, edited by Cristina Cacciari and Patrizia Tabossi, 229-248. Hillsdale: Erlbaum.

Swan, Mike. 1997. "The influence of the mother tongue on second language vocabulary acquisition and use." Mikeswan.net. Accessed 13 April 2022. https://mikeswan.net/wp-content/uploads/2017/08/The-influence-….pdf

Sökmen, A. 1997. "Current Trends in Teaching Second Language Vocabulary." In: *Vocabulary: Description, Acquisition, and Pedagogy*, edited by Norbert Schmitt and Michael McCarthy, 237-257. Cambridge: Cambridge University Press.

Uppsala Universitet. 2022. "Uppsala Student English Corpus (USE)" *Department of English*. Electronic resources. Accessed 12 April 2022. https://www.engelska.uu.se/research/english-language/electronic-resources/use/

Vaurio, Leena. 1998. "Lexical inferencing in reading in English on the secondary level." *Jyväskylä studies in education, psychology and social research* 145. https://jyx.jyu.fi/handle/123456789/13353

Wood, Mary McGee. 1986. *A Definition of Idiom*. Indiana: Indiana University Linguistics Club.

Wray, Allison & Michael R. Perkins. 2000. "The functions of formulaic language: An integrated model." *Language & Communication 20*, issue 1: 1-28.

Wulff, Stefanie. 2008. *Rethinking Idiomaticity: A Usage-based Approach*. London: Continuum.

Yle. 2017.”” En mä sitä katseella pökerryttänyt, se ehkä ulospäin näytti siltä” – Valentin
        Kononen muistetaan oksennuksesta ja kiinalaisen kumoon tuijottamisesta” [”I didn't
        make him faint by staring, it might have looked like it” – Valentin Kononen is
        remembered for throwing up and staring down a chinese person]. Yle, 22 of
        December. Accessed 22 March 2022. https://yle.fi/aihe/artikkeli/2017/12/22/en-ma-
        sita-katseella-pokerryttanyt-se-ehka-ulospain-naytti-silta-valentin

Zimmerman, C. B. 1997. "Historical trends in second language vocabulary instruction". In:
        Second Language Vocabulary Acquisition, edited by James Coady and Thomas
        Huckin, 5-19. Cambridge: Cambridge University Press.

Appendix

Appendix 1 Finnish summary

Idiomit ovat monisanaisia ilmaisuja, joiden merkitystä ei voi päätellä kirjaimellisesta merkityksestä, eli idiomi niin sanotusti "ei ole sanojensa summa". Idiomin määritelmiä on kuitenkin useita erilaisia. Idiomit rikastuttavat kielenkäyttöä, ja tekevät puheesta tai tekstistä elävämpää ja mielenkiintoisempaa. Sen vuoksi idiomien osaaminen on tärkeää vieraissa kielissä. Tutkimuksen tarkoituksena oli tarkastella englanninkielisessä ympäristössä vietetyn ajan vaikutusta oppijan idiomien käyttöön. Tutkimusaineisto kerättiin Uppsalan yliopiston opiskelijoiden englanninkielisestä korpuksesta (The Uppsala Student English Corpus). Korpus on kerätty vuosina 1999-2001 ja se sisältää monia erilaisia tekstejä, joita englannin kielen opiskelijat ovat kirjoittaneet kurssitöinä. Korpus sisältää myös opiskelijoiden taustatiedot, kuten äidinkielen, englannin kielen arvosanan ja englanninkielisessä ympäristössä (esim. vaihto-opiskelu) vietetyn ajan. Kulttuuriaiheista kurssia varten kirjoitetuista esseistä valikoitiin yhteensä 50 tekstiä: puolet opiskelijoilta, jotka olivat viettäneet aikaa englanninkielisessä ympäristössä 0-6 kuukautta, ja puolet opiskelijoilta, jotka olivat viettäneet englanninkielisessä ympäristössä yli 6 kuukautta. Tekstit luettiin tarkasti ja merkittiin ylös niistä löytyneet idiomit, jotka myös luokiteltiin alatyypin mukaan kolmeen eri kategoriaan perustuen Fernandon (1996) luokitteluun. Seuraavaksi kerron tutkimuksen teoreettisesta taustasta.

Idiomit ovat osa laajempaa kokonaisuutta, eli vieraan kielen sanaston oppimista. Aikaisemmin sanaston oppiminen nähtiin toissijaisena verrattuna kielioppiin, mutta nykyään sanaston merkitys on ymmärretty paremmin (Mäntylä 2004, 17). Sanan oppimisen eri vaiheet ja sanojen opettamisen metodit ovat saaneet erityistä huomiota. Pitempien sanaryppäiden, kuten idiomien, tärkeys sanavaraston osina on myös ymmärretty. Myös opetuksessa sanaston tärkeys on otettu huomioon ja pääpaino ei ole enää pelkän kieliopin opettelussa. Sanastotietoutta voidaan mitata monin erilaisin tavoin: suosittu tapa on mitata kuinka monta sanaa oppija osaa ja verrata sitä siihen, kuinka monta hänen tulisi osata. Erään arvion mukaan oppijan tulisi osata noin kolmetuhatta sanaperhettä (Niitemaa 2014, 142). Äidinkieli voi avustaa vieraan kielen sanaston oppimisessa, mutta siitä voi olla myös haittaa. Esimerkkinä on niin sanottu "false friends" -ilmiö, jossa jokin vieraan kielen sana on hyvin samankaltainen kuin äidinkielessä esiintyvä sana, mutta merkitys onkin täysin eri.

Kielioppi-käännösmetodi on varhainen ja tunnettu sanaston opettamisen metodi, joka perustuu etymologiaan ja sanan määrittelyyn. Koulukirjoissa opetetut sanat eivät ole sattumanvaraisia, vaan perustuvat frekvenssilistoihin, jotka koostuvat usein käytetyistä sanoista, jotka kuuluvat tiettyyn teemaan. Äidinkielen käyttö sanan määrittelyssä on eduksi, sillä tutkimukset osoittavat, että oppiminen käy nopeammin, jos sana selitetään äidinkielellä (Nation 2001, 66). Liiallinen äidinkieleen turvautuminen voi kuitenkin olla myös haitaksi: liiallinen äidinkielen käyttö luokkahuoneessa ei mahdollista samanlaista tottumista vieraan kielen käyttöön kuin pelkän vieraan kielen käyttö opetustilanteissa (Nagy 1997, 73-74). Luonnollinen lähestymistapa nojaakin vieraan kielen käyttöön: oppijat voivat arvuutella sanojen merkityksiä kontekstin avulla tai turvautua vieraskieliseen sanakirjaan (Sökmen 1997, 237). Toisto ja mahdollisuus uusien opittujen sanojen käyttämiselle ovat myös hyvin tärkeitä. Pelkkä sanoille altistuminen ei välttämättä johda tyydyttävään lopputulokseen sanavaraston laajenemisen kannalta (Vaurio 1998, 47). Edes pitkäaikainen oleskelu vieraskielisessä valtiossa ei välttämättä johda kovin suureen sanavaraston laajenemiseen, vaikka reseptiivisessä sanavarastossa on havaittu parannusta edistyneillä oppijoilla (Marton, siteerattu Mäntylän tekstissä vuodelta 2004 sivuilla 21-22). Täysin sattumanvarainen sanaston omaksuminen vaikuttaisi olevan melko epätehokasta (Sökmen 1997, 237-238), mutta riittävästi ohjattu lukeminen vieraalla kielellä voi johtaa tehokkaampaan sanaston omaksumiseen (Coady 1997, 235). Sattumanvarainen sanaston omaksuminen on yhteydessä kielitaitoon: parempi tietämys vieraasta kielestä on yhteydessä tehokkaampaan sanaston omaksumiseen.

Aiheet ja teemat ovat saaneet isomman roolin sanaston opetuksessa, kun sanojen välisten yhteyksien merkitys on tunnustettu. Sanasto tulisi siis nähdä sanojen ja merkityksien verkostona yksittäisten sanojen sijaan. Tästä ajatuksesta on kehitetty opetuksen metodi, joka nojaa sisällön opettamiseen käyttämällä pelkästään vierasta kieltä ilman äidinkielen avustusta. Etuna on se, että tällöin opetus keskittyy enemmän merkitykseen kuin muodollisiin seikkoihin (Brinton et al. 1989, 4). Lukumetodi ja tilanneoppiminen ovat yhteydessä työskentelyyn pelkällä vieraalla kielellä (Zimmerman 1997, 9-10). Vieraalla kielellä lukemista on pidetty parhaimpana tapana kasvattaa sanavarastoa (Krantz 1991, 10). Oppija joutuu havaitsemaan sanan, irrottamaan sen kontekstista, ja miettimään tai arvuuttelemaan sen merkitystä jollakin tavalla (Nation 2001, 63-65). Vieraan sanan merkitys voidaan selittää monella eri tavalla, kuten kuvilla, kontekstin avulla tai äidinkielisellä määritelmällä. Opetuksen kommunikatiiviset lähestymistavat korostavat kielenkäyttöä ja sen eri funktioita (Zimmerman

1997, 12-13) jolloin merkitys nousee tärkeäksi tavoitteeksi. Pääpaino on puhutussa kielessä ja suullisessa sujuvuudessa. Kommunikatiivisten lähestymistapojen tehokkuudesta ei ole vielä paljoa näyttöä, mutta ne ohjaavat oppijaa käyttämään kieltä aktiivisesti ja voivat tehostaa sanaston oppimista myös. Lisäksi laaja sanavarasto on hyödytön ilman kykyä hyödyntää sitä kielenkäyttötilanteissa, jolloin myös kieliopin osaaminen on tärkeää.

Sanaston opetus on siis muuttunut vuosien varrella samalla kun kielenopetuksen yleiset trendit ovat kokeneet muutoksia. Nykyään vaikuttaisi olevan niin, että erilaisia lähestymistapoja voidaan hyödyntää erilaisilla oppijoilla heidän vaihtelevien ominaisuuksiensa ja erilaisten tavoitteidensa mukaisesti. Sanastontutkimuksen monimutkaisuus heijastuu lisääntyneissä sanastoa koskevissa tutkimuksissa. Kielenopetuksessa huomiota on alettu kiinnittää myös enemmän sanaryppäisiin yksittäisten sanojen sijaan (Mäntylä 2004, 24).

Seuraavaksi kerron sanaryppäistä ja kaavamaisesta kielestä (formulaic language). Kaavamainen kieli liittyy läheisesti idiomeihin ja idiomaattisuuteen. Diskurssianalyysi, eli kielen tarkastelu syntaktista tasoa syvemmältä, koki suosionnousun 1970- ja 1980-luvuilla (Wulff 2008). Diskurssianalyyttiset metodit korostavat sitä, että fraseologiset yksiköt eivät ole marginaalinen ilmiö. Oikeastaan ne ovat hyvin yleisiä ja siis hyvin tärkeitä kielen yksiköitä. Fraseologian piirissä ei ole vakiintuneita termejä näille monisanaisille yksiköille vaan termejä on monia, esimerkiksi idiomaattiset ilmaukset (Fernando 1996), ja vakiintuneet ilmaukset (Moon 1998). Wray ja Perkins (2000, 1) ehdottavat kaavamaisen kielen olevan etukäteen muodostettu sarja sanoja, jotka varastoidaan ja haetaan muistista yhtenä ryppäänä niiden käyttöhetkellä sen sijaan, että ne pitäisi erikseen muodostaa tai analysoida kieliopin mukaisesti. Idiomit ovat yksi kaavamaisen kielen alalaji. Muita alalajeja ovat esimerkiksi kollokaatiot ja tervehdykset.

Kaavamainen kieli luo yhteyden sanaston ja kieliopin välille. *Hei mitä kuuluu?* on kysymys joka voidaan analysoida kieliopillisesti mutta se on myös yksikkö, jota kielenpuhujat osaavat käyttää kokonaisuutena analysoimatta tarkemmin sen osia. Kysymys esitetään tietyssä kontekstissa ja siihen vastataan yleensä tavalla, joka on myös etukäteen muodostettu yksikkö, jota ei tarvitse analysoida: *Hyvää kuuluu. Entä sinulle?* Cowie (1998, 1) esittää, että etukäteen muodostetut yksiköt ovat olennaisia natiivitason kielitaidon saavuttamiselle. Pitää kuitenkin muistaa, etteivät kaikki oppijat tavoittele natiivitasoa. Tietoisuus kaavamaisista ilmauksista

lisää ymmärrystä: on olennaista esimerkiksi tietää mitä tarkoitetaan ilmauksella *potkaista tyhjää*, jotta asiat voidaan ymmärtää oikein.

Seuraavaksi kerron idiomeista, jotka ovat kaavamaisen kielen alalaji. Idiomille ei löydy yhtä pätevää määritelmää, vaan määritelmiä on monia erilaisia. Termit idiomi ja idiomaattisuus liittyvät läheisesti toisiinsa, mutta eivät tarkoita samaa asiaa. Mäntylä (2006, 152) sanoo että idiomit ovat samankaltaisia kuin muut sanaryppäät, kuten kollokaatiot ja sananlaskut, ja niitä voi olla vaikea erottaa toisistaan. Mäntylä (2004, 28-35) luettelee eräitä idiomin tunnusmerkkejä: metaforisuus, analysoitavuus, muodon jähmeys, muodollisuuden taso ja monisanaisuus. Häkkinen (2002, 451) ja Nenonen (2002, 7-8) esittävät konventionaalisuuden olevan yksi idiomin tunnusmerkki, eli idiomin tulee olla yleisesti tunnettu ja käytetty kieliyhteisön sisällä. Idiomaattisuudella tarkoitetaan fraasien taipumusta saada merkityksiä, jotka ovat syvemmällä kuin fraasin yksittäisten sanojen merkitykset. Idiomaattisuus on siis kompositionalisuuden vastakohta: idiomaattisen ilmauksen yksittäisiä sanoja ei voi tarkastella yksinään ja näin päätyä oikeaan tulkintaan sen merkityksestä.

Idiomit voivat olla luonteeltaan metaforisia, mutta ne eroavat metaforista. Itseasiassa monet idiomit ovat niin sanottuja *kuolleita metaforia*. Tämä tarkoittaa sitä, että aikaisemmin käytetty metafora on menettänyt kuvainnollisen merkityksensä ja näin ilmaus on muuttunut ajan saatossa idiomiksi. Mäntylä (2004, 36) tosin toteaa, että idiomin erottaminen metaforasta ei ole välttämättä tarpeen tai edes mahdollista. Idiomeita on perinteisesti pidetty jähmeinä, eli niiden muotoa ei voi muuttaa riskeeraamatta ilmauksen idiomaattisuutta (ibid., 28). Jotkin muutokset tosin ovat mahdollisia. Jotkin idiomit pystytään esimerkiksi muuttamaan passiiviseen muotoon (ibid., 32).

Fernando (1996, 35) luokittelee idiomit kolmeen alakategoriaan: puhtaat idiomit, puoli-idiomit ja kirjaimelliset idiomit. Puhdas idiomi on täysin ei-kirjaimellinen ilmaus kuten *potkaista tyhjää*. Mikään tämän ilmauksen yksittäisistä sanoista ei ole kirjaimellinen merkitykseltään. Puoli-idiomit sisältävät sekä kirjaimellisia että kuvainnollisia osia: esimerkkinä englannin kielessä käytetty *drop names*, jossa osa *names* viittaa kirjaimellisesti sanoihin. Puoli-idiomeissa usein verbi saa merkityksensä vain esiintyessään tietyn sanan kanssa, esimerkiksi sanan *names* esiintyessä verbin *drop* kanssa. Kirjaimelliset idiomit ovat helppoja ymmärtää eli merkitys on läpinäkyvä: esimerkkinä englannin kielen *tall, dark and handsome*.

Seuraavaksi kerron idiomien oppimisesta ja opettamisesta vieraan kielen opetuksessa. Idiomien ymmärtäminen on tärkeää vieraan kielen osaamisen kannalta, sillä ihmiset harvoin käyttävät vain kirjaimellisia ilmauksia puheessaan. Idiomeita käytetään muun muassa kohteliaisuussyistä ja välittämään ajatuksia, jotka olisivat muuten hankala kuvata sanoin. Kaavamaisen kielen omaksumista ja ymmärtämistä ei ole laajasti tutkittu, vaikka se on tärkeä osa sanastoa. Tämä voi johtua siitä, että lingvistiikan alalla on muutenkin annettu hyvin vähän huomiota sanastolle (Mäntylä 2004, 81). Idiomit tuottavat vaikeuksia vieraan kielen oppijalle, koska ne ovat ilmauksia, joiden merkitys ei ole pääteltävissä yksittäisten sanojen kirjaimellisista merkityksistä. Laufer (1989, 12) havaitsi, että monet oppijat käänsivät idiomit kirjaimellisesti sanasta sanaan. Oppijoille kohdistetussa puheessa saatetaan pyrkiä välttämään idiomaattisia ilmauksia, ja oppijat törmäävät idiomeihin vain tilanteissa, joissa ei ole mahdollista neuvotella merkityksestä, kuten esimerkiksi televisiota katsoessa. Oppijat voivat käyttää monia erilaisia strategioita idiomin merkityksen päättelyyn. He voivat esimerkiksi käyttää kontekstia apuna, miettiä kirjaimellista merkitystä tai löytää äidinkielestään vastaavanlaisen idiomin (Cooper 1999). Äidinkieli voi myös vaikeuttaa vieraan kielen idiomin ymmärtämistä oikein. Riskinä ovat samalta vaikuttavat idiomit, joiden merkitys onkin täysin eri, eli niin sanotut *false friendsit*.

Seuraavaksi kerron tutkimuksen analyysista ja tuloksista. Tutkimuskysymykset olivat seuraavanlaiset:

1. Millaisia eroja voidaan havaita ryhmien välillä? Käyttääkö toinen ryhmä esimerkiksi enemmän jotakin tiettyä idiomien alalajia?

2. Onko ilmoitetulla kokemuksen määrällä englanninkielisessä ympäristössä vietetystä ajasta jonkinlainen yhteys idiomien käyttöön esseissä?

50 tutkimukseen valikoidusta esseestä löytyi yhteensä 41 idiomia. 26 esseessä idiomeja ei esiintynyt lainkaan. 0-6 kuukautta kokemusta englanninkielisessä ympäristössä oleskelusta ilmoittaneen ryhmän 11 esseetä sisälsi idiomeja, ja yli 6 kuukautta kokemusta ilmoittaneen ryhmän 13 esseetä sisälsi idiomeja. Ensiksi mainitussa ryhmässä idiomeja oli 16 kappaletta ja toiseksi mainitussa 25 kappaletta. Ensiksi mainitussa ryhmässä kirjaimellisia idiomeja oli 1, puoli-idiomeja 7 ja puhtaita idiomeja 8. Toiseksi mainitussa ryhmässä oli 4 kirjaimellista idiomia, 8 puoli-idiomia ja 12 puhdasta idiomia. Jälkimmäinen ryhmä siis käytti enemmän idiomaattisia ilmauksia. Lisäksi jälkimmäinen ryhmä käytti enemmän erityisesti puhtaita idiomeja, joiden voidaan argumentoida olevan vaikeampia. Yllättävää on, että ensimmäinen

ryhmä käytti kirjaimellisia idiomeja vain kerran, vaikka ne ovat helppotajuisempia. Tosin voidaan argumentoida, että kirjaimelliset idiomit ovat harvinaisempia muutenkin, koska suuri osa idiomeista on ainakin osittain kuvainnollisia.

Vaikka ryhmien välisiä eroja tarkastellessa vaikuttaa siltä, että enemmän kokemusta englanninkielisessä ympäristössä oleskelusta johtaa lisääntyneeseen idiomien käyttöön, yksittäisiä esseitä ja niiden kirjoittajien taustatietoja tarkastellessa ei havaittu isompien ilmoitettujen kuukausimäärien johtavan siihen, että esseessä olisi käytetty useampia idiomeja. Oikeastaan vaikuttaa siltä, että ilmoitettujen kuukausien lisääntyessä on esseissä käytetyillä idiomeilla taipumus vähentyä, vaikka voisi olettaa, että kokemus on hyväksi ja kokeneemmilla oppijoilla on parempi kielitaito joka osa-alueella ja täten myös paremmat valmiudet käyttää idiomaattisia ilmauksia. Pitää kuitenkin muistaa, että tutkimuksen aineisto oli hyvin rajallinen. Asiaa voitaisiin tulevaisuudessa tutkia tarkemmin suuremmalla aineistolla ja selvittää, löytyykö kokemuksen määrän ja idiomien käytön välinen korrelaatio tilastollisessa analyysissa.