PROCESSING OF ENGLISH ARTICLES AND THE IDIOM PRINCIPLE

A Re-examination of the Phraseological Perspective

Master's thesis

Tuomas Hartikainen

University of Jyväskylä
Department of language and communication studies
English
September 2018

JYVÄSKYLÄN YLIOPISTO

Tiedekunta – Faculty Humanistis-yhteiskuntatieteellinen tiedekunta	Laitos – Department Kieli-ja viestintätieteiden laitos
Tekijä – Author Tuomas Hartikainen	
Työn nimi – Title Processing of English Articles and the Idiom Principle: A Re-examination of the Phraseological Perspective	
Oppiaine – Subject Englannin kieli	Työn laji – Level Maisterintutkielma
Aika – Month and year Syyskuu 2018	Sivumäärä – Number of pages 70 + Liitteet 2 sivua

Tiivistelmä – Abstract

Tämän maisterintutkielman tarkoitus oli replikaatiotutkimuksen avulla tarkastella aikaisemmassa tutkimuksessa esitettyä väitettä siitä, että englannin artikkelien käyttöä voi helpottaa niin kutsuttu idiomiprinsiippi, eli se, että kielenkäyttäjillä on muistissa hallussaan suuri määrä valmiita tai puolivalmiita kielellisiä rakenteita, joita ei kieltä tuottaessaan tarvitse enää analyyttisesti purkaa kielioppiin nojaten. Lisäksi tutkimuksessa käsiteltiin suomalaisten lukiolaisten metalingvistisiä taitoja artikkelien käyttöön liittyen.

Tutkimukseen osallistui kuusi keskisuomalaista lukiolaisryhmää, joista kaksi koostui toisen vuoden, ja neljä kolmannen vuoden opiskelijoista. Määrällistä ja laadullista tutkimusdataa kerättiin opiskelijoiden englannin tunnin aikana artikkelintäyttötestin ja kyselyn avulla.

Artikkelitestissä opiskelijat lukivat tekstin, josta oli poistettu kaikki artikkelit, ja lisäsivät tekstiin oikeat artikkelit. Testin sisään oli piilotettu fraasipareja, joissa artikkelin käyttö oli saman kielioppisäännön mukainen, mutta fraasiparit erosivat toisistaan yleisyydessään englanninkielessä. Artikkelitestin jälkeen opiskelijoita pyydettiin selittämään joitakin heidän omista artikkelivalinnoistaan, ja tarkoitus oli vertailla, olivatko selitykset samat yleisten ja harvinaisten fraasiparien kesken.

Artikkelitestin tulosten perusteella opiskelijoiden oli merkittävästi helpompi syöttää tekstiin oikea artikkeli silloin, kun se esiintyi osana yleisempää fraasia harvinaisempiin fraaseihin verrattuna. Lisäksi opiskelijat perustelivat artikkelinkäyttöään yleisempien fraasien kohdalla enemmän intuitioon ja muistiin tukeutuen.

Tutkimus vahvistaa käsitystä kielen formulaisuudesta, ja että ilmiö ei rajoitu ainoastaan leksikseen, vaan myös perinteisesti kieliopin hallitsemiksi luultuihin kielellisiin rakenteisiin. Kielten opetuksessa on siten syytä pohtia, voiko formulaisuus auttaa joidenkin kielellisten rakenteiden omaksumisessa, mutta yhtä lailla on valmennettava oppijoita tuottamaan ja tulkitsemaan heille uudenlaisia rakenteita.

Asiasanat – Keywords
phraseology, article system, articles, language learning, language processing, formulaic sequence, formulaic language, metalinguistic knowledge

Säilytyspaikka – Depository JYX

Muita tietoja – Additional information

Table of Contents

List of tables

List of figures

1. Introduction

Native Finnish speakers who endeavour to acquire other languages are often faced with the challenging task of gaining purchase on different systems of articles that are idiosyncratic to a given language. These systems are used for expressing grammatical definiteness of noun phrases, which is something that does not exist in the Finno-Ugric Finnish language in a similar form as in many Indo-European languages, such as English, French, German, and the second official language of Finland, Swedish. While definitiveness or the lack thereof can certainly be expressed by different means in Finnish, Finnish lacks articles that are used for marking indefiniteness such as *a, an, un, une, en, ett* etc. and definiteness such as *the, le, la, den, det* etc.

Thus, as the vast majority of young Finnish school children (around 90% of 3rd graders; Opetushallinnon tilastopalvelu Vipunen 2016) begin their foreign language studies with English, they must slowly start accustoming to a system where noun phrases are marked with either *a, an, the* or the "zero article" in discourse, and that their usage depends on a complex system of rules and assumptions about whether the hearer is familiar with whatever is being referred to. At first this is usually addressed pedagogically by instructing a set of "rules-of-thumb" to suit the learners' cognitive level, but as they progress in their studies they will almost certainly find these rules insufficient to cover all cases of article use. Unsurprisingly no consensus exists on the best pedagogical approach from which to cover article use in the classroom, and some teachers have been cited referring to the article system as an unteachable aspect of English (Leśniewska 2017: 68).

However, as opposed to the generative account of grammar that sees language essentially as a product of constant rule application, linguists nowadays have become increasingly aware of the formulaic nature of language and language processing. For decades, formulaicity has been seen as the key to language fluency among native speakers (Pawley and Syder 1983), and consequently this has had an effect on classroom pedagogy among non-native speakers as well (Meunier 2012). The fact that language users seem to take advantage of a vast resource of pre-existing sequences of lexis that are relatively effortlessly retrieved from the long-term memory is a phenomenon recognised decades ago (Wray 2002: 7-8), but especially with better access to extensive corpus data in electric form researchers have been fully able to unravel the underlying formulaicity in languages. One of the most cited ideas in this field is undoubtedly Sinclair's

(1991) *open-choice principle* and *idiom principle*, the latter of which is central to this paper and will be introduced in section two.

Much of the research on formulaic language has so far focused on formulaicity as a lexical phenomenon, concentrating, for example, on idioms and other non-compositional formulaic items that have been shown to have an advantage in language processing over novel, "nonconventional" stretches of language. To broaden this view to integrate lexis and grammar, Leśniewska (2016) presented an empirical study in a non-native setting regarding article processing from a phraseological point of view, examining the effect that formulaicity had on the processing of articles appearing in frequently and rarely occurring sequences. Based on her results she concluded her paper with an intriguing argument suggesting that article use is not only a rule-governed task, but instead there exist certain psycholinguistic mechanisms related to formulaicity that can facilitate their use.

For such a conclusion to be accepted, similar results must obviously be achievable among different populations of non-native English learners. The purpose of the present paper, therefore, is to first validate and then expand on the argument presented by Leśniewska and attempt to confirm her claim that the choice of articles by non-native students of English, whose first language is articleless, is to some extent relied on or facilitated by the idiom principle, in other words, that processing of holistic units of language negates the need for online grammar application. To fulfil this purpose, a small article-filling test – using the same instrument as in Leśniewska's (2016, appendix) with some modifications – was conducted among Finnish learners of English (N=113). To complement this data, the same participants were asked to shortly motivate some of their article choices in order to find out about their metalinguistic knowledge and whether they were consistent in their responses between items of low and high frequency that had the same underlying grammar.

The outline for this paper is as follows: first, it will briefly delve into the English article system approaching it from a cognitive point of view illustrating the difficulties learners face if and when they attempt to apply and process the rules of article use. Next, an overview of research on the formulaicity of language is provided, and the distinction between speaker-external and speaker-internal formulaicity relevant to this study is introduced. The fourth chapter describes the research proceedings conducted for this paper, including the research questions and data gathering methods. The data is then analysed in the fifth chapter, and finally, discussed in the sixth and final chapter.

2. English Article System: A Stumbling Block Even for the Advanced Learner

The article system of English is undoubtedly one of the most difficult aspects of English grammar, often said to be among the final stumbling blocks for even the most proficient non-native speakers after they have acquired all other aspects of the language. And considering the fact that even native speakers of English do not always agree on article use and interpretations of noun phrases (Butler 2002: 475), the difficulty for non-natives is not surprising. DeKeyser (2005: 5) notes how articles "express highly abstract notions that are extremely hard to infer, implicitly, or explicitly, from the input", and as such the article system makes for an interesting topic of research into cognitive processing. Plenty of research has indeed been committed over the article system and its acquisition both in first language and non-native speaker contexts, only a fraction of which shall be reviewed below.

In this paper, the following categorization of articles and their use for reference is made after Quirk et. al. (1985: 265-288):

Definite article	**Indefinite article**	**Indefinite plural and mass**
the	a(n)	*zero article*

The discussion is therefore limited to these three overt articles, while other central determiners that are sometimes used in an article-like manner (such as *some, any, no* as in *no book*) (Radden and Dirven 2007: 92) and which may sometimes function as pronouns (e.g. "Here's *some* for you") while the overt articles cannot (Quirk et. al. 1985: 254-255).

Other alternative descriptive accounts of article use and the semantics of articles have also been proposed. For example, Berezowski (2009) provides a history of the "zero article" in descriptions of English grammar and goes to great lengths at dispelling the "myth" surrounding its existence. He considers several instances where the two overt articles are inadmissible – proper names, predicate nominals and prepositional phrases just to name a few – and argues that such gaps are merely the result of incomplete article grammarization, and that they do not form any particular sets of linguistic environments that descriptive grammarians could coherently spell out (Berezowski 2009: 46-53). For the purpose of this paper, however, the term "zero article" has been chosen to refer to instances where the absence of an article is a

significant grammatical marker, following Downing and Locke's (2002: 417) statement that it is "a category in its own right."

A comprehensive discussion of the English article and reference system is impossible to conduct in this paper, but a brief explanation of different types of nouns and ways of reference in English in contrast with the Finnish language is provided. The main source for the following analysis and examples is Radden and Dirven's (2007) book on cognitive English grammar whose approach on grammar can be considered more pragmatic and functional than some more descriptive grammars such as Quirk et al.'s (1985) in that it is "usage-based" (Radden and Dirven 2007: XI) and looks at the lexicon and constructions in a language as a set of choices for a language user to select the appropriate ways for communication.

Although to some second language (L2) learners articles might seemingly have a minor and unimportant role, and while some language instructors may place emphasis on communicative competence over metalinguistic knowledge, articles nevertheless are the most frequently occurring function words in English (Master 1997) and they have an important role in communication and negotiation of meaning. A cognitive viewpoint is therefore beneficial when the object of research are participants whose native language has no articles, because their way of thinking about grammatical reference might be different.

2.1. Types of Nouns and Reference in English

In discourse, we are constantly making references to various instances of *things* (Radden and Dirven 2007: 41-57; Langacker 1991): objects and substances which have different inherent properties that affect how we understand them as ontological beings, and also as such determine how we refer to them. First, they may have perceivable boundaries like a *car* or no inherent boundedness like *water*, and their internal composition may change if they are broken down into smaller pieces, like the *car* when it is taken to a scrapyard, whereas the identity of a homogenous subject, such as *sand* or *dust* does not change even upon dimensional manipulation. While several entities of the same subset of objects can be added up or duplicated, subjects cannot, as there would emerge no individual countable elements of the same subject even when divided into portions. However, as examples such as "*Beer* tastes good – I'll have three *beers*, please!" and "I caught *a fish* – We'll have *fish* for dinner!" illustrate, some nouns have a hybrid ability of behaving like both countable and non-countable objects and subjects. With some

collective nouns encompassing multiple individuals, such as *jury* and *team* one may – particularly in British English – highlight their individual members or the group as a whole via verb agreement. In addition, in the category of concrete objects, some are regarded as intrinsically plural (*pluralia tantum*) and they are expressed as plurals in form but can require either singular or plural verb agreement ("*The news* is real"; "Our *wages* are low") (Radden and Dirven 2007: 63-78).

In contrast to these concrete things and their linguistic counterparts expressed as concrete nouns, a large number of all the things in our discourse refer to abstract things, which Radden and Dirven (2007: 78-83) describe as episodic situations or states, and steady situations or states. Unlike concrete nouns that are grounded in the physical domain, abstract things are often perceived as relations which go through a conceptual shift allowing us to refer to them as if they had ontological existence (*marriage*, for example, from the relation of *being married*). Similar to concrete things, they can be encoded either as objects or substances (count nouns and mass nouns) depending on whether they are seen as episodic (such as *attack* or *idea*) or continuous (such as *information*, *happiness*) states or events. However, there is considerable overlap between the categories, as examples such as "*War* is hell – *Wars* fought in the 20th century" show.

It is a prerequisite for successful communication that both the speaker and the hearer agree on which instances of things (**referents**) are being referred to (in a communicative act of **reference)**, and as such communication always involves pragmatic negotiation of how these instances are established in the minds of the discourse participants. The speaker may use different **referring expressions** – noun phrases – in order to ground these instances to the hearer's mental space. This includes making several assumptions based on the speaker's knowledge and the hearer's assumed knowledge of all the possible instances of the thing that is referred to (Radden and Dirven 2007: 87-89). There are various expressions used for grounding, but the following discussion is limited to how we refer by using the articles.

Radden and Dirven (2007) make a distinction between two types of reference: individuative and generic reference, which differs somewhat from Quirk et. al.'s (1985: 265) distinction between specific and generic reference. The concept at this level, however, is the same: individuative or specific reference focuses on an individual specimen of a class of entities, whereas generic reference refers to the class as a whole. The differences arise when they describe the individuative and specific references further. In Radden and Dirven's more

cognitive approach, individuative reference is divided into indefinite and definite reference, and indefinite reference further into specific and non-specific reference. According to them, the difference between specific and non-specific indefinite reference is that specific reference refers to a factually existing entity in the speaker's, but not the hearer's mind. A non-specific referent, on the contrary, exists only virtually. In an example such as "I bought *a car*" the reference is specific, because *a car* does exist in the speaker's mind, whereas in "I need *a car*", car, as of time of speaking exists outside of reality (Radden and Dirven 2007: 88-112).

The referring expressions for indefinite specific and non-specific references, however, are the same (Radden and Dirven 2007: 94-96, 111; Downing and Locke 2006: 418), which is perhaps the reason why descriptive grammars, such as Quirk et. al. (1985) do not regard them as separate. Besides, both of the above examples illustrate the use of the indefinite article *a*/*an* equally: it is used when the referent is not mutually identifiable by both the hearer and the speaker and must therefore be first instanced in the mind of the hearer for further possible elaboration.

Definite reference, on the contrary, is used when the referent(s) can be mentally shared by both the speaker and the hearer, either by its uniqueness or by general knowledge of the world. Radden and Dirven (2007: 95-105) identify three types of definite reference: deictic reference, discourse reference and unique reference. First, deictic ("showing, pointing") reference refers to referents that can be accessed and pointed out in the environment in the immediate situation where the discourse takes place. Several types of determiners are used for this reference (*this*, *that*, *here*, *the same time*/*place* etc.).

Second, discourse reference includes two types of reference that are made possible as a discourse progresses: anaphoric reference is used to point to something that was mentioned earlier in the discourse, which can be done directly by mentioning the referent again, or indirectly by using the hearer's general knowledge about the referent so that they may infer their relationship, such as in the example "John bought *a bicycle*, but when he rode it one of *the wheels* came off." (Quirk et. al. 1985: 267). In cataphoric reference, Radden and Dirven (2007: 99) explain, the referent is referred in advance as definite for it to be introduced immediately afterwards, for example when announcing a topic that the speaker will follow on. Quirk et. al. (1985: 268-268) also describe the cataphoric use of the definite article when it is followed by a postmodification that uniquely defines the referent, such as "The president *of Finland*".

Finally, referents can be unique within the socio-cultural boundaries that both the speaker and hearer share, which makes them identifiable. Radden and Dirven (2007: 99-105) identify three

types of unique reference. First, some mass nouns and proper names have inherent uniqueness, since the former examples of which include abstract nouns like *life, society, education* – represent notions known by all members of the discourse community and thus lack the need for pragmatic introduction, and the latter often points to a single instance without involving a category. Names of countries and geographical areas, on the other hand, are not as simple, as some take the definite article and some do not. Here the conceptual factor of the boundedness of the referent may help, as, for example, most articleless country names and geographical names refer to entities whose boundaries can be perceived, such as in the case of countries, cities, lakes and mountains, whereas proper names that take the definite article – names in plural, rivers, mountain ranges to name a few – are often less easily perceived as single entities (Radden and Dirven 2007: 100-101). Still, the rules (if such can be established in the first place) for the use of articles in proper names are complex, and all of them cannot be discussed here. The reader is referred to e.g. Quirk et. al. (1985: 288-297) for a broader descriptive account of proper names and article use.

Radden and Dirven's second and third type of definite references are qualified uniqueness and framed uniqueness, which correspond to instances that are made unique through descriptive linguistic expressions that isolate the referent from its class of other referents (e.g. "My dog is the *one with fluffy hair*"), or which become unique upon activation of a shared conceptual frame in the immediate or wider socio-cultural speech situation (e.g. "*The roses* are very beautiful" said in a garden; "*The murderer* left his fingerprints on the knife" said during a crime investigation.)

There is only one type of reference left to summarize: generic reference. As mentioned above, when we make a generic reference we focus on a whole class of instances instead of an individual, and like individuative reference it can be definite or indefinite. According to Radden and Dirven (2007: 107), no language has separate determiners for expressing generic reference, and in English both the definite *the* and indefinite *a(n)*, as well as the zero article are used in generic reference. Like individuative reference, generic reference can also be indefinite or definite and either singular or plural, resulting in four possible expressions of generic reference. Although in some cases very similar in meaning and thus interchangeable (e.g. "*A tiger* hunts by night", "*Tigers* hunt by night" and "*The tiger* hunts by night"), the different references do have some differences in use, which are minor enough to be omitted in the present discussion (but see for example Radden and Dirven 2007: 107-112).

All the types of reference can be summarized in the following table that shows all the types of definite and indefinite individuative and generic references.

Table 1 Main types of reference (adapted from Radden and Dirven 2007: 111).

reference								
individuative					generic			
indefinite		definite			indefinite		definite	
specific	non-specific	deictic	anaphoric	unique	singular	plural	singular	plural
I bought **a house**	I want **a house**	Look at **this house!**	Those houses; **they** look spacious	Open **the door!**	The life of **a lion**	**Girls** are strong	**The lion** hunts in packs	We should help **the poor**

The purpose of the above summary, brief and incomplete as it may be, of various types of noun phrases, types of reference and article use in English has been, from a cognitive point of view, to illustrate the myriad of choices a speaker must constantly make during a speech situation in order to create new frames of reference and utilize existing ones for the benefit of the hearer and mutual understanding. That is, if such decision making was necessary in the first place. For natives, this process becomes highly automatic relatively early in childhood, but for a language learner, acquiring and using this system can prove challenging, especially if the linguistic means of reference in the target language differ considerably from their L1 – such as the Finnish students in this paper. The difficulties acquiring the English article system have been well documented, but before delving into the literature on article acquisition studies, let us first briefly look at reference in the Finnish language.

2.2. Reference and (in)definiteness in Finnish

The above discussion of nouns and reference has been limited to the English language, but all the concepts – concrete, abstract and mass nouns, their inherent properties, as well as different reference types – can also be found on a conceptual level in Finnish as well, although there may be differences in terminology (see Hakulinen et. al. 2004: 547-556, 1349-1352). Of greater interest is how linguistic features, such as the lack of articles, affect how different references are realized in Finnish.

Hakulinen et. al. (2004: 1349-1362) explain that in Finnish there are several means for signalling definiteness or indefiniteness besides extralinguistic means (context, gestures, etc.).

Some specifiers – *eräs, muuan* (“some” or “a certain”); *yksi* (“one”) – can be used to tell the hearer that familiarity with the referent is not assumed. *Joku ~ jokin* (“some[one/-body]”) as an indefinite specifier implies that the speaker himself is unaware of the identity of the referent, or indifferent about it; it is not necessarily an individuative type of reference. In spoken language proadjective forms *semmoinen, tämmöinen, tuommoinen* (“that/this kind [of]”) can also be used to mark indefiniteness, and even in some cases the demonstrative pronouns *tämä* and *tuo* (“this”, “that”) can sometimes be used in this manner. An indefinite specifier can also be used with a person name to single out one person among many people with the same name, or similarly to English usage of *a* or *some* in “There’s a John Smith to see you”, where the speaker signals that even though they know the person’s name, they cannot fully identify them (Hakulinen et. al. 2004: 1353-1355).

Specifiers that point to definiteness include demonstrative pronouns (although note above), adjectival modification that restricts the uniqueness to a certain entity, and the pronoun *se* (“it”) (Hakulinen et. al. 2004: 1356). *Se* pronoun, as a matter of fact, when used as a specifier of definiteness, seems to be so common in spoken Finnish that there has been some debate over whether *se* is undergoing grammaticalization into a definite article (Hakulinen et. al. 2004: 1359; Juvonen 2000; Larjavaara 2001). For example, Larjavaara (2001) acknowledges the “article like” usage of *se* in some situations, but states that it cannot be considered an article due to its non-compulsory nature and specific function in spoken language.

When it comes to bare, unspecified noun phrases in spoken language, they can be interpreted as either generic references. In written Finnish definiteness is less commonly marked with specifiers, and thus they can be interpreted as definite, indefinite, or open in this regard. Their interpretation, nevertheless, can be guided with word order (Hakulinen et. al. 2004: 1360). In addition, the Finnish case system allows speakers to signal quantitative (in)definiteness (e.g. ”Puuhun tuli *omenia*[plural partitive case]”, ”*Apples* grew on a/the tree”; ”*Omenat*[plural nominal case] putosivat puusta”, ”*The apples* fell from a/the tree”) (Hakulinen et. al. 2004: 1361-1363).

As can be seen, languages with articles express different references and definiteness or the lack thereof especially in a very different manner than languages with no article system which instead rely more on specifiers, determiners, word order and the case system, like in Finnish language. Dissimilarities between reference systems are often credited for a part of the

difficulties learners face when trying to acquire a system in another language (Harb 2014). The following brief review of literature will look at these issues.

2.3. Difficulties in L2 English Article Acquisition

Plenty of research has been committed on the English system of articles, which has been aptly described as "an example of an interface phenomenon cutting across the domains of morphosyntax, semantics, and pragmatics" (Zdorenko and Paradis 2011: 39). Indeed, such a cognitively demanding system, as we saw in the previous chapters, can provide interesting insights in language learning in both L1 and L2 contexts. In native context, it has been concluded that children learning English as their native language seem to acquire the article system gradually and almost effortlessly with high accuracy (although they tend to overuse the definite article in those cases where the referent is unknown to the hearer) by the time they reach age four (Butler 2002: 454). This is in stark contrast with non-native speakers, whose difficulty in mastering the English article system has been well documented by research (Butler 2002; Herranen 1977; Vartiainen 1979; Zdorenko and Paradis 2011). In some cases, even advanced learners have been reported as unable of reaching native-like accuracy in article use (White 2003).

These difficulties can pertain to factors relating to the lexico-syntactic structure of English and discoursal elements that were discussed above – noun countability, definiteness and specificity among others (see also Harb 2014) – but the learner's mother tongue can also help or hinder the acquisition by having a similar (e.g. Spanish, French), semi-similar (e.g. Arabic) or dissimilar system (e.g. Finnish, Chinese) (Harb 2014: 98-99). Especially learners whose L1 conveys reference in a significantly different manner from English seem to be at a disadvantage in article acquisition.

For example, Snape, García-Mayo and Gürel (2012) compared Spanish (N=50), Turkish (N=88) and Japanese (N=33) ESL learners and their performance in a forced choice elicitation task where they had to select the correct article for different types of generic noun phrases in the context of a short conversation. They were interested in the transfer effect from the participants' first languages, which either had both definite and indefinite articles (Spanish), only the indefinite article (Turkish; although not an article per se, *bir* ("one") precedes nouns), or no articles at all (Japanese), and which all expressed genericness on a noun phrase level and

sentence level differently from English. In addition, the researchers predicted the sort of errors that would likely surface as a result of these differences. Based on their results they claim, for example, that ”if the L2 learner group experiences problems with article choice it is directly related to L1 transfer effects” (Snape et. al. 2012: 20). For instance, while the Spanish seemed to benefit from their L1 in selecting the definite article for the appropriate noun phrases, the Turkish and Japanese had more problems with them. Furthermore, if the Turkish could benefit from the indefinite article-like *bir* to some extent in indefinite generics, the Japanese, lacking any L1 aid, performed worse. Bare plurals at the sentence level also showed some evidence of L1 transfer, but at the same time the participants made errors with mass nouns that could not have been a result of L1 effect (Snape et. al. 2012: 22). Thus, some errors seem to be explainable by L1 negative transfer, while some not.

The relationship between metalinguistic knowledge and L2 competency has also received attention in studies on article acquisition. These studies have usually attempted to measure the explicit metalinguistic knowledge (Ellis 2009) of participants and then compared the amount to their performance in using articles. Butler, for example, (2002) studied Japanese students (N=80) of varying proficiencies with an article filling test followed by an interview with the researcher in which the participants were asked to provide the reason(s) for each of their article choices. Unsurprisingly the test scores clearly increased with more proficient students, but there was also a significant gap between the most proficient group and the control group of native English speakers.

Analysis of the participants’ metalinguistic knowledge showed that they had major problems in two particular areas. Regardless of proficiency, firstly, the participants tended to misdetect specific reference or hearer knowledge (or both), or they failed to consider referentiality altogether, and secondly, they were susceptible to misdetection of noun countability. Butler proposes that these two obstacles are intertwined in the sense that in order to detect hearer knowledge of a referent, one must first be able to identify whether the referent is countable and therefore belongs to a set of referents which might or might not be identifiable to the hearer (Butler 2002: 473-474). Differences in countability and boundedness of noun phrases has also been reported as a source of difficulty among, for example, Korean students (Amuzie and Spinner 2013). Analysing the explanations further, Butler found that less proficient participants were more prone to using nongeneralizable or idiosyncratic hypotheses in their answers compared to the more advanced students who were more often successful in identifying the correct reason for article use. Similar results on the relationship between explicit metalinguistic

knowledge and language performance have been reported elsewhere (Elder 2009: 115-117), but there still remain important issues on how exactly explicit and implicit learning interface in relation to each other (Ellis 2009: 20-23).

Finnish learners of English and their article use in particular has been studied to a small extent, although much of the available work is already decades old. Herranen's (1977) study examined compositions written by university students (N=90) from different English levels and furthermore employed a multiple-choice article test among first-year university students of English (N=45). Vartiainen (1979) also studied article errors in compositions written by comprehensive school students in the sixth, seventh, eighth and ninth grades (aged 12-15 years). In more recent work, Lehtonen (2015) analyzed and compared the use of English and Swedish articles (the article systems in the two languages share many similarities; see Lehtonen 2015: 28-45) in compositions written by university students, which were then rated according to the Common European Framework of Reference.

Vartiainen found that in written products the use of the indefinite article in different contexts caused most difficulty for among all her participants, correct usage being less than 50% in all grades and omission being the most used strategy, especially if the noun had adjectival premodification (Vartiainen 1979: 86-87). Although the definite article was mastered better, omission was also frequent, but here the presence of premodiciation seemed not to affect accuracy. The learners had not yet mastered the use of articles in generic reference, according to Vartiainen (1979: 90-91), which resulted in low accuracy in generic references. Herranen's (1977) both analyses from the compositions and the multiple-choice test, in contrast, showed that uses of the definite article seemed to cause the most problems, accounting for over half of student errors and suggesting that specific reference was more difficult than non-specific. Generic reference caused also major problems (Herranen 1977: 48-49). Both Herranen and Vartiainen noticed a tendency to omit the article if the noun phrase was premodified with an adjectival attribute, which has also been observed by Trenkic (2007) among Serbian students. Lehtonen's (2015: 113-114, 119) hierarchy of difficulty also shows that it was the various categories of definite references that university students made most mistakes both in English and Swedish, reaching around 80-90% accuracy. Students on all CEFR levels seemed to use generic and indefinite references relatively well, and Lehtonen (2015: 216) concludes that for university students the zero article seems the easiest, followed by the indefinite article and the definite article.

In conclusion, while uses of the indefinite article seemed to be more difficult for younger students while the definite article was mastered better, it is important to note that in Vartiainen's (1979) data of student compositions, for example, unique references rarely occurred in more demanding contexts than, for example, *the sun* or *the sky* and that mass and abstract nouns only occurred in positions that required the zero article (Vartiainen 1979: 90-91). Therefore, no conclusions could be made concerning the pupils' command of these reference categories. Meanwhile both papers by Herranen (1977) and Lehtonen (2015) papers with more advanced participants and higher frequencies of each reference category (albeit unevenly distributed and in some cases few in number as well) pointed to definite and generic references causing the most problems among Finnish learners.

What is noticeable in all the studies discussed above is that they all seem to implicitly consider article usage essentially as a grammatical, rule-governed and cognitive task. And while there is debate over the exact role of explicit and implicit metalinguistic knowledge in language production (Ellis 2009) and what pedagogical practices for teaching the article system are the most effective, most research seems to agree that some form of instruction is necessary (for example, Akakura 2012, Master 1997, Master 2002). L2 learners are therefore expected to have acquired *some* explicit linguistic knowledge via formal instruction and they are, often in the context of article acquisition research, seen as using this knowledge or exhibiting the lack of it (as for example in Butler 2002).

However, the following chapter will present a different view on how language, including the use of articles, can be processed. Instead of seeing language users as constant appliers of explicit or implicit rules of language, this view highlights the formulaic nature of language and that a large part of our language use actually consists of prefabricated sequences that are stored and retrieved from the long-term memory, thus facilitating both language input and output processing. Leśniewska (2016: 217) calls this "the phraseological perspective".

3. Formulaic Language: The Key to Fluency

Pawley and Syder (1983) were puzzled by the fact that native speakers are able to convey their intended meaning by using expressions that are consistently both grammatical and idiomatic even though there are almost limitless amounts of other possible yet more or less unidiomatic utterances a speaker could make to convey the same thing. They give an example of an utterance a host of a party could speak to a friend who arrives with a mutual friend: "I'm so glad you could bring Harry!", which is a perfectly natural and unmarked phrase, as opposed to many other, much less ordinary but grammatically fine utterances that exhibit the speaker's command of the language, such as "That Harry could be brought by you makes me so glad", "That you could bring Harry gladdens me so", "Your having been able to Harry bring makes me so glad" and so forth (Pawley and Syder 1983: 195-196). Pawley and Syder's answer to the puzzle was that nativelike command of language relies heavily on knowledge of a large number of lexicalized sentence stems that can be edited appropriately for different contexts.

This and other similar observations (for example Sinclair 1991) have since given spark to a great deal of research in different fields of linguistics on such *formulaic* approach to language, which is sometimes contrasted with the analytical view that language is generated via grammar. During the last decades, especially computer-based corpus linguistics and increased access to vast amounts of corpus data have given rise to endeavours into phraseological elements in language (Cowie 1998, Granger and Meunier 2008, Moon 1998, Sinclair 1991, Wood 2010, Wray 2002; see also Wray 2013 for an excellent timeline review of research in formulaic language). It was Lamb (1998: 169) who stated that "Linguists seem to underestimate the great capacity of the human mind to remember things while overestimating the extent to which humans process information by complex processes of calculation rather than by simply using prefabricated units from memory", and indeed, by now it has become well established that the Chomskian idea that language production among adult native speakers rests upon their ability to construct utterances with the power of analytical grammar is not enough to explain how language is processed. Instead, native speakers and L2 learners alike, to a very great extent, take advantage of various pre-existing, holistically stored sequences of language, such as idioms, idiomatic expressions, fixed or semi-fixed sequences and collocations during language processing (Wray 2002). This is largely considered to be the key to fluency among non-native speakers as well, as Wray (2000: 463) states: "Gaining full command of a new language

requires the learner to become sensitive to the native speakers' preferences for certain sequences of words over others that might appear just as possible." Formulaicity, then, is ubiquitious in language use (Conklin and Schmitt 2012: 46), and estimations of how large a portion of discourse – written and spoken – consists of formulaic language vary according to the methods of analysis. For example, Erman and Warren (2000) found that prefabricated language accounted for over 50 percent of native-level spoken and written text, according to their criteria of "prefabs" (discussed in section 3.1.).

A pioneering researcher in phraseology and corpus linguistics was John Sinclair whose 1991 publication *Corpus, concordance, collocation* (Sinclair 1991) outlined contemporary studies in computational linguistics illustrating the power of corpus analysis in revealing underlying lexical patterns in language, including collocations, colligations, semantic preferences and semantic prosodies (Sinclair 1991, 2004). The book also features Sinclair's perhaps most influential idea of how language is processed according to two opposing systems: the *open choice principle* and the *idiom principle*. The former is a way of looking at language text as a result of a number of choices: each syntactic position within linguistic units (words, phrases, clauses) must be filled from the lexicon while adhering to grammatical rules of the language (Sinclair 1991: 109). On the other hand, the idiom principle, similarly to Pawley and Syder (1983), suggests that "a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments" (Sinclair 1991: 110).

Taking these ideas of language processing into account, the question of interest in this paper is how the idiom principle can facilitate language processing when it comes to the use of English articles. Accordingly, some relevant studies on the processing of formulaic language will be introduced in the following chapters. Particular attention is given to Leśniewska's (2016) and Takahashi's (1997) studies, both of which are concerned on the processing of articles and the formulaicity of language. Their results seem to imply that correct article usage is to some extent facilitated by formulaicity effect and the idiom principle, and that in some cases there may not be a need for a speaker to consider all the complex rules associated with article use that were discussed above. This paper is an attempt at replicating Leśniewska's (2016) results from Polish university students among another population of ESL learners and provide more evidence for this view. But first, let us look closer at language and formulaicity and its significance in language processing.

3.1. What's in a Sequence?

As mentioned, the formulaic nature of language has now for decades been the object of formal linguistic, corpus-linguistic, pragmatic, and psycholinguistic research (Myles and Cordier 2017: 4) among native speakers, non-native speakers, adults, children, and aphasic patients alike (Wray 2013). This research, however, has been all but uniform in terms of methodology and terminology. Wray (2002: 9), for example, lists over 50 terms that have been used to describe formulaicity and its aspects in the context of the above various disciplines. To give a few examples, Erman and Warren's (2000: 31) prefabs mentioned above, were defined as combinations of two or more words "favored by native speakers" over other possible combinations "which could have been equivalent had there been no conventionalization". Such definition relies heavily on the researchers' intuition (Wray 2002: 20), which may even be considered unscientific by some, or problematic at least. Moon (1998) defines *Formulaic Expressions including Idioms* (FEIs) according to three variables of institutionalization, lexicogrammatical fixedness and non-compositionality, while intentionally excluding some phraseologically interesting units: compound words, phrasal verbs, foreign phrases and inflections of multi-word forms like "had been lying" and "more careful(ly)" (Moon 1998: 2-9). In addition to Moon's criteria, frequency count based on a corpus analysis is also often used as a criterion for detecting formulaicity (Wray 2002: 25-31).

It should not be mistaken that it were meaningful to dichotomously label expressions as "formulaic" or "non-formulaic", even if researchers such as Sinclair (1991) talk of two modes of language processing. Ellis' review (2012) shows how important a factor frequency is in the processing of multi-word sequences, and just as there is no dichotomy of "frequent" and "infrequent" but rather a continuum between these ends, so too it is more meaningful to consider formulaicity a continuum. Wray (2012) discusses the possibility of a two-way continuum with regards to frequency and compositionality: frequent and infrequent idioms and other noncompositional strings (e.g. *Osama Bin Laden*; *kith and kin*) being at one end and compositional strings (e.g. *at the end of*; *at the home of*) at the other.

Thanks to Wray's (2002) contribution to formulaic language research, the term *formulaic sequence* has become the most widely adapted term by researchers (Myles and Cordier 2017: 9) to describe psycholinguistic holistically stored multi-word units. Her definition of a formulaic sequence attempts to be as inclusive as possible in stating the following:

> "a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar" (Wray 2002: 9)

It is worth noting that Wray later called this only a stipulative definition that worked as a basis for her analysis; it was not an end product of any analysis nor something that describes and explains the phenomenon of formulaicity (Wray 2009: 29). For a discussion of the latter sort of definition and more detailed discussion of formulaicity which Wray describes as "morpheme equivalence", there is not enough space in the present paper, but instead Wray's other work is recommended as a starting point (Wray 2002: 265-269, 2009: 29-34). It is worth noting that this paper does not deal so much with the identification of formulaic sequences, but how they are processed.

Nonetheless, as mentioned, the above definition for a formulaic sequence has been deemed sufficient working term by many studying phraseological elements. Problems, however, arise when researchers use the same terminology to refer to different phenomena in linguistic and psycholinguistic contexts. Myles and Cordier (2017) argue that this has been taking place for some time now, and that there is a great need for further elaboration in the definitions used. They state that there are essentially two kinds of formulaic sequences: learner external and learner internal, which will be discussed in the next section.

3.2. Speaker-External and Speaker-Internal Formulaicity

Myles and Cordier (2017: 5) allege that in the literature researchers are using the term "formulaic sequence" on the one hand to refer to native and non-native speakers' use of idioms, idiomatic expressions and collocative sequences in a given language, and on the other hand to sequences that are idiosyncratic to an individual. The essential difference is that the former approach describes the formulaicity of a given language based on evidence that exists outside the learner, while the latter examines psycholinguistic units within a learner that are retrieved more effortlessly from the long-term memory than other strings, thus facilitating their processing. While these constructs may overlap particularly as regards to L1 speakers, that formulaic sequences derived from a corpus based on their high frequency would manifest as psycholinguistic units within any speaker's brain has been proven to be a false assumption (Schmitt, Grandage, Adolphs 2004; cf. Ellis, Frey and Jalkanen 2009). Miles and Cordier,

therefore, stress the importance of differentiating between speaker-external and speaker-internal formulaic sequences, termed thus by Wray (2008). A learner-external formulaic sequence is most likely to be psycholinguistically valid learner-internal FS in the case of a native speaker, but if the same FS is produced with errors or with difficulty by a non-native speaker, it cannot be regarded as a psycholinguistic unit retrieved holistically from the memory (Myles and Cordier 2017: 5).

Myles and Cordier proceed to coin yet two more terms to reflect the difference between speaker-external and speaker-internal formulaic sequences: linguistic clusters and processing units. Their respective definitions state as follows (Myles and Cordier 2017: 12):

> [Linguistic clusters are] multimorphemic clusters which are either semantically or syntactically irregular, or whose frequent co-occurrence gives them a privileged status in a given language as a conventional way of expressing something.
>
> [A processing unit is] a multiword semantic/functional unit that presents a processing advantage for a given speaker, either because it is stored whole in their lexicon or because it is highly automatised.

As a means for identifying processing units in L2 learners, Myles and Cordier (2017: 17-22) suggest a hierarchical method that considers phonological fluency, holistic quality, and frequency of sequences that are regarded as candidates for processing units in learner language production. In other words, sequences must be pronounced coherently without pauses, they must have semantic or functional unity (e.g. expressions that fulfil a certain purpose such as referring to time or place) or have been learned as holistic units (such as classroom routines), and the more frequently they are used the more reliable their status as a processing unit is. However, since this paper deals in learner production in a very limited sense, such method of analysis is irrelevant to this study.

Nonetheless, as the present study is conducted in a L2 context where the overlap between speaker-external and speaker-internal formulaic language is not as obvious, Myles and Cordier's recommendation about speaker external and internal formulaicity is acknowledged and referred to in subsequent sections of this paper. The next section will briefly look at literature related to formulaicity related language processing advantage, which is also relevant to this study.

3.3. Formulaic Language and Processing Advantage

According to Wray (2002: 93-102), formulaic sequences have several functions in a speech situation that aid both the speaker and the hearer, all of which can be considered to work to the speaker's interests. They may alleviate the effort for language processing for the speaker, and in some cases for language decoding for the hearer by, for example, linguistically plotting the course of discourse or signalling with a simple, commonly shared utterance ("excuse me", for instance) that the speaker wants the hearer to do something. This increases the chances of the speaker to be listened to and understood correctly. Some sequences may also belong to a certain register and using them can signal the speaker's identity. Wray (2002: 101) furthermore points out that formulaic sequences do not come from a static storage in the memory, but they are a dynamic resource that language users change according to their needs. Of these functions of formulaic sequences, the facilitating effect of in language processing is the main point of interest in the present study.

A considerable amount of research has been committed on the processing of formulaic language, and plenty of evidence has been found to support Sinclair's (1991) claim that of the two language processing principles discussed above, the idiom principle is the default one. In practice this means that processing of formulaic language is generally quicker and "potentially" different in some ways from nonformulaic language (Conklin and Schmitt 2012: 47), most likely due to the fact they are processed as holistic units (Schmitt and Underwood 2004: 173). Much of this research on processing of formulaic language has focused on idioms (for example Underwood, Schmitt and Galpin 2004; Conklin and Schmitt 2008; Siyanova-Chanturia, Conklin and Schmitt 2011), which, as Conklin and Schmitt (2012: 50) point out, can be problematic: idioms are relatively infrequent, have varying degrees of transparency (compare *clear as day* and *kick the bucket*), and can be ambiguous having figurative or literal meaning, all of which can affect their processing, especially among non-natives who may not be as exposed to idioms as native speakers.

Fortunately, nonidiomatic language has been studied extensively as well. For example, Tremblay and Baayen (2010) measured the processing of four-word strings using immediate free-recall method with native English speakers (N=11). They gathered both behavioural and electrophysical ERP data, and better recall was shown to be affected positively by the frequency of occurrence of the four-word sequence. An eye-tracking study by Siyanovia-Chanturia, Conklin and van Heuven (2011) similarly found that their native and proficient non-native

participants (N=28) read frequent three-word binominal sequences (e.g. *bride and groom*) faster than infrequent sequences, and that their reversed counterparts (eg. *groom and bride*) were read slower.

Ellis and Simpson-Vlach (2009) and Ellis, Simpson-Vlach and Maynard (2008) tested natives and non-natives (varying numbers of participants) in four experimental procedures in order to determine how corpus linguistics metrics affect accuracy and fluency in processing academic formulas (e.g. *in other words*, *in the case of the*, *in the context of the*). The items were sampled by their length, frequency and mutual information (MI) factor, which is a statistical tool for assessing the coherence of a sequence, that is, the strength of association between words. The experiments measured the speed of reading and recognition in a grammaticality judgement task, rate of vocal reading, the speed of which the final word of a sequence is read aloud when it is first primed by what comes before it, and the speed of comprehension and acceptance when the formulae were placed in either an appropriate or inappropriate context. After analyzing the results, the researchers concluded that the corpus-derived formulae did have psycholinguistic validity, and that higher frequency and MI value positively affected the processing of the formulaic sequences. Interestingly, their results suggested that in the case of natives, the MI value affected processing more than frequency, but vice versa in the case of non-natives.

In conclusion, the processing advantage of formulaic language is well grounded, but there are still issues left to be solved. For example, it is still open to debate whether the processing advantage reflects holistic processing of sequences, as Trembley and Baayen (2010) suggest, or faster mapping of individual components (Wray 2012: 233-234). In any case, the research mentioned here is only a small fraction of the available literature due to space limitations, but for example Conklin and Schmitt's (2012) review provides a more comprehensive look at recent research into formulaic language including both idioms and nonidiomatic language.

3.4. Use of English Articles and the Phraseological Perspective

We finally turn to the primary topic of this study: use of English articles in light of the formulaic nature of language, a perspective that has not been fully explored in article acquisition studies at all. Some related studies are worth mentioning, though. Leńko-Szymańska's (2012) exploratory corpus-based study, for one, set out to investigate the extent of which Polish English learners' article use could be accounted for by "conventionalized language". She measured

frequencies of 3-grams containing the definite and indefinite articles (*n*-grams are reoccurring multiword lexical bundles, as defined by Biber et. al. 1999) from Polish learner corpora consisting of compositions from different proficiency levels, and then compared the frequencies to those found in native speaker corpora (of different genres of published texts). In native corpus, 3-grams accounted for 29% and 17% of all instances of the uses of definite and indefinite articles respectively. Frequencies in the learner corpora showed that Polish students tended to increasingly rely on conventionalized language (in other words, 3-grams including articles such as *a lot of*, *there is a*, *it was the*, *to be the* etc.) as their language abilities grew, and that eventually at advanced level the frequencies surpassed those of natives (35% for uses of *the* and 23% for uses of *a/an*) (Lenko-Szymańska 2012: 11). Lenko-Szymańska further observed the frequencies of article use standardized to the size of the corpora and found that the use of articles in conventional instances reached native-like frequency with regards to the definite article and exceeded native-like frequency in the case of the indefinite article, whereas frequencies of rule-based uses remained on a much lower level even on advanced level (Lenko-Szymańska 2012: 12). This seemed to indicate an overreliance on conventionalized language in the use of the indefinite article, and overall underuse of articles in rule-based contexts.

Lenko-Szymańska's study did have a major shortcoming in that it did not consider the accuracy of the *n*-gram tokens and only analysed raw frequencies. It did not consider the use of the zero article in lexical bundles either. Nevertheless, the paper still shows how learners, as their proficiency grows, become sensitive to formulaicity in article use and that they increasingly utilize reoccurring sequences with articles (Lenko-Szymańska 2012: 16). However, whether this is due to some phraseological effect relating to the processing advantage of such sequences cannot be stated based on solely this evidence.

Some empirical studies on article use that have employed fill-in-the-article type tests have included a category of test items consisting of idiomatic expressions. Ekiert's (2004) and Li and Yang's (2010) papers in Polish and Chinese speaking settings respectively showed the difficulties their participants (N=25; N=80 respectively) had with idioms and fixed expressions such as *live hand to mouth*, *all of a sudden* and *in the face of*. Only the most proficient group of Chinese speakers reached accuracy of over 85% compared to two lower level groups (around 20% and 40%) (Li and Yang (2010: 23), while the Polish students on average reached around 50% accuracy in these items (Ekiert 2004: 14). Whereas these two studies concluded that articles within idioms and fixed phrases are especially problematic for students, Lenko-Szymańska's (2012) showed how large a proportion of learner article use actually consisted of

conventionalized use. This, as Lenko-Szymańska (2012: 15) suggests, is most likely due to different definitions of idiomatic use. The 3-grams in Lenko-Szymańska (2012) represented sequences of high frequency and the idiomatic expressions in Ekiert (2004) and Li and Yang (2010) infrequent compositional expressions.

Takahashi (1997) was perhaps the first to, in an article filling test, compare the accuracy between commonly and non-commonly occurring sequences, as he termed them. These particular items in his test included four common and non-common items listed below in Table 2, which, according to a collocation analysis (Takahashi 1997: 7), can be seen occurring most frequently with the definite article, or the indefinite article in the case of *there is* X, which can be considered to give them a status of linguistic clusters.

Table 2 "Commonly occurring sequences" and "Non-commonly occurring sequences" in Takahashi's (1997) test instrument.

Commonly occurring sequences:	
where's the coffee?	*the first* word
the third floor	*the only* person
Non-commonly occurring sequences:	
there is Ø glass everywhere	won *Ø first* prize
swimming *in a* beautiful sea off Greece	he is *a second-class* player

Takahashi hypothesized that for his Japanese university students (N=99) knowledge of the above commonly occurring sequences might, on the one hand, lead them to the correct answer in the common sequences but, on the other hand, to the incorrect answer, that is, the insertion of *the* (or *a/an*), in regard to the noncommon sequences. His results seemed to imply that this was indeed the case: the accuracy in common sequences was 53% compared to 41% in non-common sequences on a whole-group level. Among the top 30 performers these figures were 63% and 44% respectively, and among the bottom 30 performers 44% and 33% respectively. Overall results from the test including other categories of article use was 54%, suggesting that articles proved a very difficult challenge for the Japanese participants.

These results seem to be perfectly in line with those of Lenko-Szymańska's (2012) in that both Japanese and Polish learners exhibited a clear sensitivity to frequencies in their article processing. The Polish learner corpora showed how learners' article use became increasingly conventionalized as their level of English grew, and the Japanese reached higher accuracy when

they had to fill in an article in a commonly occurring sequence compared to instances in rule-based usage.

There are some issues in Takahashi's methodology, though. First, the items categorized as common or noncommon are not directly comparable because they differ syntactically and semantically. Second, the phrase *first prize* with or without the definite article is often heard in English discourse even among natives. In the Corpus of Contemporary American English (Davies, 2008; COCA) it occurs a total of 301 times – 37 times with the definite article – and in the British National Corpus (2007; BNC), its total frequency is 288 – 76 times with the definite article. The longer phrase *won* (*the*) *first prize*, however, is much rarer with the definite article, which may imply a special idiomatic status for the phrase (*someone*) *won first prize* (particularly in the past tense), which the participants may not have been sensitive to. Third, Takahashi did not disclose the frequencies of *the*, *a/an* and *zero* answers, so it is impossible to know whether the participants were led astray in their answers due to the effect of formulaicity (for example, providing the answer "he is *the* second-class player") or whether they for other reasons failed in their choice. The results would have been more reliable had Takahashi used compositional test items that were similar in syntax and semantics but differed in frequency.

All the studies discussed above present interesting contributions to the topic of article processing and formulaicity, but there remains a significant gap: in such instances of article use which occur in a compositional and rule-governed environment, does high frequency have a facilitating effect for processing of the articles, as has been demonstrated with other types of formulaic sequences?

Leśniewska's (2016) recent paper attempted to fill this gap. Her study targeted Polish university students (N=90) who were studying on either B2 (N=44) or C1 (N=46) level of the Common European Framework of Reference, and it investigated the same phenomenon as Takahashi (1997) but with greater emphasis on the lexical combinations that the target articles appeared in. She also used target item pairs that matched syntactically, in other words, the rule governing the article use was the same. Her article filling test contained 12 target pairs of frequently and rarely occurring compositional lexical combinations (see Table 3 below) hidden inside sentences of which all articles had been removed and where there were no clozes to indicate the possibility of article insertion. This was to make the task more authentic in terms of article use. The test was designed so that the effect of the vocabulary would be minimized; according

to the researcher, all participants were familiar with the vocabulary involved, so it was assumed there was no interference from difficult words or phrases (Leśniewska 2016: 212).

The test items were initially selected based on researcher intuition, and then subjected to native speaker judgement, just as Read and Nation (2004: 29-30) recommend for inter-rater reliability. The items were then analysed with the help of corpora, showing that the frequent items did indeed occur significantly more often in the corpus data than the rare combinations. Two corpora – the BNC and COCA mentioned above – were used for this analysis. Leśniewska's target items are reproduced here as a full list of frequently and rarely occurring item pairs, as the empirical study in the present paper takes advantage of this existing test, albeit with slight modifications that will be discussed in the methods section.

Table 3 Leśniewska's (2016, appendix) frequently and rarely occurring target item pairs.

Frequent target item	**Infrequent target item**
a friend of mine	*an acquaintance of mine*
what a shame	*what a remarkable player*
twice a day	*five times a semester*
the sooner the better	*the smaller the pot, the more critical the problem*
a cup of tea	*a spoonful of syrup*
the day I die	*the food I brought*
help the poor	*open to the insured*
hit (someone) in the face	*cut in the hand*
speak English	*learn Kurdish*
get a job	*live in a luxury apartment*
have kids	*eat carbohydrates*
the centre of attention	*the ecology of waterways*

The results from the Polish students showed that the mean score for frequent items (0.85) was significantly ($p<.05$) higher compared to rarer items (0.68) on a whole group level, as well as among both lower (0.75 and 0.53 respectively) and higher level groups (0.94 and 0.82 respectively). Difference in mean scores was more notable among lower level students compared to higher level students (Leśniewska 2016: 213).

A detailed look at the test item results showed that in 9 of the 12 target pairs the more frequent combination had a significantly ($p<.05$) higher score than the infrequent item. In one pair – *the centre of attention* and *the ecology of waterways* – the mean score difference was nonsignificant

(0.61 and 0.57 respectively), and there were two pairs of which the more infrequent item actually had a very slightly and insignificantly higher mean score: *what a remarkable player* (0.81), as opposed to *what a shame* (0.79), and *live in a luxury apartment* (0.90), as opposed to *get a job* (0.89) (Leśniewska 2016: 215). The latter item pairs will be discussed in more detail in the methods section of this paper, as they received some modifications in the present study.

Regardless of the above three item pairs, that the overall results exhibited much higher accuracy for the more frequent items, Leśniewska argues, is evidence for her claim that was mentioned in the introduction and repeated here word by word: ”the perception of article use (outside of idiomatic uses) as being purely rule-governed may be incomplete and should be broadened to include - - the phraseological perspective” (Leśniewska 2016: 217). This perspective essentially holds that the psycholinguistic mechanisms related to sensitivity and more fluent processing of formulaic language extend to aspects that have been traditionally regarded as analytical and rule-governed processes. In practice, more frequent and thus more formulaic combinations of noun phrases are likely to be processed with the idiom principle, rather than by online grammar application, which is in line with all previous research on the processing of formulaic language. But whereas previous studies have mainly considered the processing advantage of formulaic sequences to be a lexical phenomenon, Leśniewska’s approach broadens this view to integrate the lexis and grammar.

4. The Present Study

This chapter introduces and discusses the empirical study conducted in this paper. The purpose was to gather both quantitative and qualitative data from a specific group of ESL learners: Finnish gymnasium (upper secondary school) students in the later stages of their studies. The research questions and their respective hypotheses are introduced first, followed by a description of the participants and test methodology. Special attention is paid to the differences in methodology between the present study and that of Leśniewska's (2016) so that the reader gets a clear idea of the extent to which the present paper is a replication of the original.

4.1. Research questions

The research questions for this study were as follows.

1. Does the idiom principle account to some extent for the correct article use by learners of English (Leśniewska, 2016: 209)? If yes, what is this extent among students who receive lower and higher English grades?
2. In which target items can the facilitating effect of the idiom principle be discerned?
3. What sort of metalinguistic knowledge do the participants portray in order to motivate some of their article choices? Are their justifications consistent between items that are syntactically similar but which differ in formulaicity?
4. What kinds of misconceptions about the article system do the participants portray in their grammar explanations?

First and foremost, the main purpose of this paper was to find further evidence for the claim made by Leśniewska (2016) that there is a certain facilitating factor in article processing when they appear as a part of a more frequent and thus more formulaic expression. The second aim was to explore this facilitating effect further by scrutinizing the differences between proficiency groups and the target items. The a priori hypothesis behind the first and the second research question was that the idiom principle does indeed facilitate the choice of articles in formulaic sequences even among Finnish students, and that this effect is more noticeable in less advanced

learners, because the capability of article processing in both rare and frequent occurrences naturally increases with one's overall language proficiency, as Leśniewska (2016: 217) mentions. It was also expected that the effect of formulaicity would not necessarily be observed in all target item pairs, as, for example, item pairs such as *speak English* and *learn Kurdish* might prove relatively easy for the participants so that significant differences would not occur.

The third and original aim of this paper was to find out how the learners motivate some of their article selections and whether they could verbally explain their metalinguistic knowledge – an analysis not present in Leśniewska's (2016) paper, but something that for example Butler (2002) conducted among Japanese learners of English. The target items the participants had to provide metalinguistic explanations for consisted of four pairs of high and low frequency clusters which had the same underlying grammar. Of particular interest was to see whether the participants were consistent in their explanations or whether they would tend to treat either of the combinations differently. For example, perhaps in the case of the frequent combinations the students would tend to rely more on their intuition, whereas in the case of rarer sequences they might have to think of the explicit grammar rule given the hypothesis that the non-formulaic sequences would have to be processed according to the open choice principle, and thus, in a more discreet manner. These questions are worthy of exploration as they may lead into new insights into an aspect of English that is particularly challenging to Finnish learners and may contribute to improvement of teaching practices.

4.2. Participants

Six groups of Finnish ESL students (N=113) from three different gymnasium schools located in Central Finland participated in this study in spring 2018. With all the participating groups data gathering took place at the beginning or the end of one of their regular English lessons of which about 30 minutes were spend on completing the study, each student proceeding at their own pace. The test instrument and instructions were printed on A4-size paper and the students used pencils to fill them. Two groups consisted of second year students (N=29; 16-17 years of age) in their fifth gymnasium English course, while four groups were formed by students in their third and final year (N=84; 17-18 years of age) and who were on a preparation course for the Finnish matriculation examination in English language when the data gathering took place. The test was advertised as an opportunity to revise the use of English articles, and all participants were told that after completing the test they would have the possibility to see the

correct answers, which would be followed by a brief introduction to the research topic given by the researcher. This was hoped to become an incentive for the students to participate in the study to the best of their abilities.

Finnish ESL learners were chosen as the target group of research due to the absence of grammatical articles in the Finnish language. As mentioned above, research has shown that native speakers of languages without articles seem to be at the most disadvantageous position when acquiring a foreign language with an article system, and that Finnish students of English even up to the university level may have difficulties with the article system. It was therefore expected that the Finnish target group would face a reasonable challenge with the test instrument used in this study, which in turn was hypothesized to make the facilitating effect of formulaicity of language to emerge.

4.3. Methodology and Test Design

For the present study, the same test instrument as in Leśniewska (2016, see section 3.3. above for a description of the test) was used, but with some target item modifications which are discussed below in the following chapter. The students were instructed to read each test sentence carefully once or twice and insert what they felt were the correct missing articles. As in the original version, no gaps were present to indicate the possible choice of an article. This part of the test will be referred to as the test phase.

Whereas Leśniewska (2016: 212) states that all participants in her study were familiar with the vocabulary used in the test instrument, the same claim cannot be made in the present study due to the fact that the participating students came from three different schools with different study backgrounds. Nevertheless, the vocabulary used in the test was deemed suitable for the gymnasium students by both the researcher and the instructor of this paper. Furthermore, during the test phase some of the rarer words and their translation into Finnish were provided for the students by writing them visibly on the classroom blackboard. These words were *carbohydrates* (*hiilihydraatit*), *waterway* (*vesireitti*), and *insurance policy* (*vakuutussopimus*).

To complement this data, a written rule verbalization task was designed to elicit explicit metalinguistic knowledge from the students, following the example of Hu (2011). The definition for metalinguistic knowledge is borrowed from Purpura (2004: 88), who terms it as "informational structures related to linguistic terminology". After completing the article filling

test, the students proceeded to the backside of the test instrument where they were asked to refer back to their own article choices in eight specific locations of the test phase and in their own words motivate these particular choices of articles – including instances where they had not selected any articles. Students were asked to use Finnish in their explanations so that no interference would result from lack of competence in English. For example, the first item of this phase read in Finnish:

a. Look at your article choice in sentence 2: …**whenever we eat ___ carbohydrates.** How do you motivate your article choice in the underlined part?

This phase of the study will be later referred to as the reasoning phase. The test phase and the reasoning phase were kept separate by instructing the students not to look at the backside of the test instrument before they had completed the article filling test completely, and after turning to the backside they were not allowed to change any of their article choices. Thus, it can be assumed that they paid no particular attention to the eight locations specific to the reasoning phase while they were completing the article filling test.

In the reasoning phase, all participants were encouraged to think about the grammatical rule first when writing their answers, but they were also told that answers such as ”I don’t know” and ”It just sounded good” were acceptable answers as well. The aim was, on the one hand, to avoid too many ”I don’t know” replies, but on the other hand also to avoid too many responses left blank. Due to the nature of this study, even answers such as ”This sounded good to me” were considered valuable, because they at least gave an indication of some intuition related psycholinguistic mechanism at work.

The four target item pairs of which their article choices the participants were asked to motivate were *have kids* and *eat carbohydrates; the sooner the better* and *the smaller the pot the more critical the problem; the day I die* and *the food I brought; a friend of mine* and *an acquaintance of mine*. The choice to include only these items while excluding 16 other target items was due to practical limitations of the research setting. As the study was conducted as a part of a regular English lesson, no more than a moderate amount of time could be required to complete it. Moreover, if asked to write several motivations more, the participants might have become demotivated and less likely to provide meaningful answers.

The above eight items were chosen to represent four cases of article use: generic indefinite reference (*eat carbohydrates, have kids*), individuative definite reference (*the food I brought,*

the day I die), individuative indefinite reference (*an acquaintance of mine, a friend of mine*) and a grammatical construction (the comparative correlative) formed with the definite article (*the sooner the better, the smaller the pot the more critical the problem*). Out of Leśniewska's (2016) target items, these items were the ones in their respective grammatical categories which had the highest divergence in their mean scores. Similar variances in the mean scores were expected in the present study as well, and it was considered interesting to see whether the students could nevertheless express the same grammatical rule for both item pairs or not.

This method of data gathering was chosen because it allowed for a large enough population of respondents in order to make statistical analysis and generalizable conclusions from the test phase of the study (Taanila 2014). It also allowed for gathering of both qualitative and quantitative data. Of course, there are some limitations in the present methodology. Particularly the method for gathering data on how the participants portrayed their metalinguistic knowledge is considerably more limited than, for example, Butler's (2002) proceedings; she had an approximately 30-minute-long interview with each of her participant concerning all one hundred target items in her test instrument, which resulted in much more elaborate responses. In contrast, in the present study one blank line on an A4-size paper was provided for a student's written response. The space was therefore limited and answers were expected to be very short, offering a limited range of grammatical hypotheses, which is consequently reflected in the concise nature of their analysis.

4.4. Test Modifications from Leśniewska (2016)

While discussing the results of her study, Leśniewska inspects three target item pairs in which the effect of formulaicity was not perceived (see section 3.3). To repeat, these three items were *what a shame* and *what a remarkable player*; *get a job* and *live in a luxury apartment*; *the centre of attention* and *the ecology of waterways*. Regarding the first item pair, Leśniewska (2016: 214) points out two things: the difference in the degree of countability of the nouns shame and player, as well as the adjectival premodification in *remarkable player* may have affected the processing of these two items. It has been shown that for some learners accuracy with article use depends on the boundedness of the noun phrase (Amuzie and Spinner 2013), as well as that learners tend to omit articles more often in items that are adjectivally premodified (Herranen 1977, Trenkic 2007, Vartiainen 1979), although the latter did not seem to occur in Leśniewska's (2016) study. One point I would like to add, albeit a simple one, is that in Leśniewska's test

isntrument these two target items occurred in very close vicinity to one another, separated by only one short sentence. The particular test item read: ”16. What remarkable player he is. His performance today really impressed me. What shame he didn't get picked for team.” (Leśniewska 2016: appendix) It is possible that their close vicinity resulted in unintended saliency for the items, and that participants became aware of the need of an article in both items if they noticed at least one of them.

Due to these observations and their possible implications, instead of the item *what a remarkable player* another item, *what an accomplishment* was used. The adjectival premodification was removed, and the fact that only ten instances of the phrase were found in the COCA and none in the BNC, it was deemed sufficient to label it as a ”low frequency item”. Although the issue of the difference in degree of countability and boundedness is still present, it was considered an experiment to see whether the higher formulaicity of *what a shame* would in this case prove easier for the students. The two items were also separated from one another in order to reduce their saliency. They can be found in sentences 12 and 16 in the test instrument (see appendix A).

As for the second item pair, *get a job* and *live in a luxury apartment*, Leśniewska notes that the latter item was, despite it being an infrequent item, rather easy for her participants, as an apartment is a countable, concrete noun. She suspects that due to ”some sort of a ceiling effect” (Leśniewska 2016: 214) the effect of formulaicity failed to register in this item pair. Thus, the infrequent item *live in a luxury apartment* was changed into *get a wage increase*, which appeared only once in the COCA and was not found at all in the BNC.

The third item pair that did not exhibit significant difference was *the centre of attention* and *the ecology of waterways*. As there seems to be no plausible explanation for this, this item pair was not altered in the present study in order to confirm the result and to account for random variability. The complete test instrument with the above modifications can be found in Appendix A.

4.5. Coding

The SPSS statistics software (version 21) was used for all the statistical analyses of the test scores in this paper. Quantitative data for the 24 target items and the students’ reasons for their article choices were input into the program in several forms: first, each participant received

either one or zero points for each individual correct or incorrect use of article in the 24 target items. If the indefinite article was required, participants were awarded a point as long as they wrote either *a* or *an*, because the aim was to see whether the learner was able to correctly choose between the definite, indefinite and zero-article, and not between the two indefinite articles. Individual points were summed into variables representing the sums of correct high and low frequency items, as well as the total correct sum. Of these sums means were then calculated to represent the scores with a number between 0 and 1.

Then, the participants' motivations for eight of their article choices were coded into one of several categories. The encoding process was completed by the present researcher, and while this does raise issues concerning subjective judgement during the encoding process, all answers were processed in a succession under a specific guideline in order to minimalize errors due to rater subjectivity. What follows is a description of this process.

4.6. Coding of the Qualitative Data

For this part of the analysis, data from nine participants was excluded due to more than half of their answers having been left completely blank in the reasoning phase. This was assumed to have resulted from the time running out and thus proper consideration of the items was not likely to have taken place. All these nine students were from the lower proficiency group. Answers from the remaining 104 participants were encoded into six distinctive categories, categorization being a useful tool for handling large amounts of qualitative data and exploring patterns inside it (Rob 2004).

The first one included explanations that were judged to reasonably demonstrate the participant's understanding of the article use context. It is important to note that this did not mean that the student had to be able to perfectly lay down the underlying grammar rule in words, but rather in some way demonstrate his or her understanding of concepts such as definiteness or the lack thereof. Therefore, instead of "correct", these answers will be referred to as "adequate" or "acceptable" in the analysis section. How each target item pair was handled and what was considered adequate is explained further below after explanation of the other reply categories.

The second category included answers where the student attempted to explain the logic in his or her article use but failed to meet the criteria for an adequate reasoning (again, see below). These answers were thus labelled "inadequate". The qualitative analysis in section 5.2.2. will

concentrate on this response category. The analysis was carried out by identifying typical patterns of student responses and then identifying the reason they were regarded as inadequate, that is, whether they were incomplete or portrayed some sort of misconceptions about the English article system.

In the third category were included answers that were either left blank or which did not state any cognitive reason at all for the participant's article choice, or which only mentioned the difference between the use of indefinite articles *a* and *an*. Thus, answers such as "I don't know", "I guessed", "I didn't put anything here" and also "The word acquaintance begins with *a*, so I put *an*" were placed under this category. The latter were included in this category and not the second one because they failed to consider why the definite article was necessary in the first place, which was the object of this inquiry. Such explanations will be referred to as "unattempted".

The fourth category, to make a distinction between the previous one, included answers that referred to the participant's intuition. In Finland, this intuitive grasp of correct language is often referred to as *kielikorva* (literally "language ear") and the rather large number of answers mentioning *kielikorva* showed that for many students this is a valid strategy for motivating their article (and very likely other grammatical) choices. As mentioned above, it was considered important to categorize these answers separate from the answers in the third category, because they can be seen as an indication of some psycholinguistic mechanism at work, whereas "I don't know" answers could be the result of mere guesswork. Due to the methodology, it is, unfortunately, impossible to accurately define whether this actually is the case. There is a possibility that the students used the *kielikorva* explanation when they had simply guessed the answer. However, several test sheets showed how the same student had written both "I don't know" and "This sounded good" type answers, which implies that they had at least thought of them as separate types of answers. This category's label is "intuition" in the analysis below.

Interestingly, some students also referred to the phrase-like quality of the target item using one of the Finnish words for "phrase", "saying" or "proverb". Sometimes they even mentioned that they had heard the phrase in, for example, a movie or a song. These answers clearly suggested how some of the target items were indeed regarded as holistic units and that the students assumedly used this information to their advantage in the test phase. To see which items were regarded as such, explanations referring to this sort of phenomenon were placed under the fifth category, with the label "phrasal".

Finally, some participants wrote that they did not understand some parts of the target item and were thus unsure about their article choice. While few in numbers, these instances clashed against the presupposition that the vocabulary in the test played no role in the results, and thus such answers were deemed important to categorize separately. In the analysis, they are referred to as "unintelligible".

There were several instances where the participant had written some form of grammatical explanations for their choice while also mentioning that it sounded good by "the language ear". In such overlapping cases, the grammatical explanation was considered to be the decisive element, that is, whether it was adequate or not, and the answer was categorized accordingly. Also, when a participant referred to another one of their answers (e.g. "same as above"), the response that was referred to was regarded as the explanation for that item as well.

What sort of explanation was judged to be acceptable depended, of course, on the target item pair. First, for *eat carbohydrates* and *have kids* the participants were expected to be able to explain that carbohydrates and kids were referring to a plural and generic (i.e. non-definite) entity. Explanations that mentioned only plurality were considered inadequate, because both words could take the definite article in plural form in some other contexts. Second, in the case of *the sooner the better* and *the smaller the pot the more critical the problem*, explanations that mentioned the presence of a grammar rule or which named the Finnish equivalent of *the – the* construct (*mitä – sitä*) were accepted. Third, in *the day I die* and *the food I brought* some leniency had to be applied, because while many participants correctly explained that either *the day* or *the food* in question was specific, they did not mention what made them specific in these contexts. The choice was made to accept explanations that only mentioned the specificness of these items even without reference to the postmodifying part, which actually made these items specific. Lastly, both *a friend of mine* and *an acquaintance of mine*, on the opposite end, referred to non-specific entities, so explanations that used some word choice to refer to their non-specificity or the absence of specificity were accepted as adequate.

5. Results and Analysis

This chapter will report on the data gathered for this paper, which was both qualitative and quantitative in nature. In the first part, mean scores for high and low frequency items and all items in total are calculated and examined, and then the target item pairs are investigated individually. For the second part, data from the participants' reasons for their article choices are analysed both quantitatively and qualitatively.

For the purpose of analysis, participants were divided into two groups of higher and lower proficiency based on the calculated average grades from their three latest English gymnasium courses that the participants themselves reported. Naturally, an objective "proficiency level" for a student cannot be determined solely by using his or her school grades, because grades – a product of summative assessment – seldom tell the whole story about a student's real-world language competence. But since there was little possibility to conduct a more robust assessment of the students' language skills due to lack of time, the school grades, assumed to reflect their overall language competences, were used.

Before this, however, an independent sample t-test, which can be used for comparing average scores between two populations (Faherty 2008: 215-220), was run in order to determine whether there were significant differences in the mean test scores between the two groups of second-year students and the four groups of third-year students, which would have suggested that the scores were not comparable between these groups due to the fact that the third-year students had been studying English for a longer time. The t-test, however, showed that although mean test scores were slightly higher for the groups of third-year students, these differences were non-significant ($p>0.05$). This was deemed sufficient evidence to say that the article filling test score was not significantly dependent on whether the participant was a second-year or a third-year student. The average school grade, on the other hand, seemed to be a better predictor of a participant's test score among both second and third year students: Pearson's *r* correlation test, which is better suited for scaled data than Spearman's rank correlation (Faherty 2008: 183-190), showed that a significantly high correlation (.752, $p<.001$) existed between higher English grades and higher overall test scores. Therefore, students with an average grade of equal or less than 8.0 (the Finnish grading scale is from 4 to 10) were placed into the lower group (N=63) and the rest into higher group (N=50).

Five students did not report their English grades for some reason, but they could nevertheless be placed into either group based on their article filling test score. A simple linear regression analysis, which is often used for examining the relationship between two or more variables (Berk 2004), was conducted to predict a participants' grade based on their test score. The analysis yielded a significant regression equation ($F(1,106)= 138,045$, $p<.001$) with R^2 of .562, and the following formula was used to calculate a dummy grade for these five students: $a = 4.944 + 4.650b$, where *a* is the student's grade and *b* the mean overall test score. Of course, the possibility that these students were placed into "wrong" proficiency groups exists, as their actual English grades may have suggested a different placement. Due to there being only five such students whose overall test scores were clearly either higher or lower than the average score, this was considered to be a non-critical issue.

5.1. Mean Scores of Frequent and Rare Combinations

Results similar to Leśniewska's (2016) study could be obtained in the present study when the mean scores for frequent and rare combination items were calculated and significant differences in them examined. Paired t-test was utilized for this analysis, as it is a powerful tool for comparing two mean scores and finding out whether the difference is statistically significant (Faherty 2008: 201-207). The numbers from this analysis are provided in Table 4 below. Although the mean scores of all items combined are of relatively little interest from a phraseological point of view, they are still included for the sake of completeness and also as an indicator of the students' general article processing abilities. Here the data shows that these participants were, on average, accurate only 62% of the time in their article choices for the target items (excluding, of course, other instances of article use that were inside the test instrument).

Table 4 Mean test scores of frequent and rare combination items

	All items (*SD*)	Frequent (*SD*)	Rare (*SD*)	Paired t-test (Frequent-Rare)
All participants (N=113)	0.62 *(0.20)*	0.69 (*0.21*)	0.56 (*0.23*)	t = 9.57, p < **.001**
Higher proficiency (N=50)	0.77 *(0.13)*	0.83 (*0.11*)	0.71 (*0.17*)	t = 6.38, p < **.001**
Lower proficiency (N=63)	0.51 *(0.18)*	0.58 (*0.20*)	0.44 (*0.20*)	t = 7. 13, p < **.001**

As can be seen in Table 4, the difference between the mean scores of frequent and rare combination items was 0.13 on a whole-group level, which is similar, but somewhat smaller than that of Leśniewska: 0.17 (Leśniewska 2016: 213). This means that in both Polish and Finnish groups the same phenomenon was observed: the participants tended to choose the correct article more often when the choice appeared in a more frequent lexical combination, such as *a friend of mine*, as opposed to those instances of articles that were part of a rarer combination inside the test instrument, for example *an acquaintance of mine*. Both groups of higher and lower proficiency groups displayed this phenomenon, but between the frequent and rare items there was not as large a gap between the groups as in Leśniewska's data; the gap between frequent and rare sequences was 0.12 for the more proficient students and 0.22 for the less proficient group, whereas in the present data the mean differences were almost equal (0.12 and 0.14 respectively).

Of course, these mean scores are not one hundred percent comparable due to the fact that the test instrument in these two studies was different. However, the changes made in the present paper concerned only 2 of the 24 target items, which does not change the overall trend in the mean score analysis, as will be seen in the following section.

5.2. Analysis of Individual Test Items

Analysing each combination pair individually, one can yet again perceive a similar outcome as in Leśniewska's (2016) study: not all pairs of frequent and rare word combinations seemed to exhibit the phenomenon that article use was easier in the context of the more frequent item. The

mean scores, standard deviations and statistical significance of the difference in means calculated with a t-test are presented in Table 5 below. The items are presented in the same order as in Leśniewska's paper for ease of comparison, version signifying whether the item is a frequent (A) or rare (B) combination. Bracketed words are included to list the items exactly as they appeared inside the test instrument.

Table 5 Mean test scores of item pairs

Version	Article	Target item	*Mean*	*t*	*p(t)*
A	a	*a friend of mine*	.61	6.337	**<.001**
B		*an acquaintance of mine*	.27		
A	a	*what a shame*	.67	-1.521	.131
B		*what an accomplishment*	.74		
A	a	*twice a day*	.58	.928	.335
B		*five times a semester*	.53		
A	the	*the sooner the better*	.65	6.771	**<.001**
B		*the smaller the pot the more critical the problem*	.31		
A	a	*a cup of tea*	.89	1.807	.073
B		*a spoonful of [this] syrup*	.82		
A	the	*the day I die*	.94	4.993	**<.001**
B		*the food I brought*	.73		
A	the	*help the poor*	.32	3.440	**.001**
B		*open to the insured*	.16		
A	the	*hit [me] in the face*	.40	1.450	.150
B		*cut in the hand*	.31		
A	zero	*speak English*	.96	.446	.657
B		*learn Kurdish*	.95		
A	a	*get a job*	.88	6.237	**<.001**
B		*get a wage increase*	.56		
A	zero	*have kids*	.94	1.748	.083
B		*eat carbohydrates*	.88		
A	the	*the centre of attention*	.43	-1.221	.225
B		*the ecology of waterways*	.50		

As can be seen from Table 5, the present study found five target pairs of which the more formulaic combination was significantly ($p<=.001$) easier for the participants, whereas in Leśniewska's (2016) study significant differences were found in nine of the twelve pairs. It is natural that different populations would yield different results due to different backgrounds of English study. For example, how much instruction in article use they have received and whether it was in recent memory would naturally affect the results. But the overall trend seems clear, as seen in Table 4 above, and despite the fact that the present study found several cases where the differences in mean scores were not significant, in ten of the twelve target pairs the mean score

for the frequent combination was higher than that of the rare combination. If similar results were to be obtained from a larger population, it is possible that some of these items, such as *have kids – eat carbohydrates* would show statistical significance as well.

Data drawn from a small population is also susceptible to even small amounts of deviation. When each participant group's performance was examined individually, it was found that one group (of second year students, N=17) showed some rather anomalous results in a couple of items compared to the other groups. For some reason, students in this group tended to choose the correct article much more often for the rare combinations *a spoonful of syrup* and *cut in the hand* compared to their frequent counterparts *a cup of tea* and *hit in the face*. Due to this, the same paired t-test was run once more but without the results from this group, which consequently made the mean differences in *a cup of tea* (.93), *a spoonful of syrup* (.82); *hit in the face* (.43) and *cut in the hand* (.27) statistically significant at $p<.05$ level. This data exclusion also resulted in the mean difference in *what a shame* (.68) and *what an accomplishment* (.78) becoming significant at $p<.05$ level, but otherwise there appeared no major changes from the results reported on a whole-group level in Table 5. Being unfamiliar with the participants in this particular group and their English study background, it is difficult to think of an explanation for why their results in the former two item pairs differed so much from the other groups.

One must obviously be careful when reporting research data in above manner; it is easy for a researcher to exclude or include particular sets of data in order to display results that appear more favourable to his thesis. The purpose of the above analysis was not to manipulate the data, but to show that even a small amount of inconsistency may skew the results in such a relatively small sample, but presumably abate in a larger one. One may, of course, argue that these are not random inconsistencies but rather repeating patterns, the truthfulness of which shall be left for future research to determine. Some further observations regarding the target items will be made below.

Two test items in which the mean score for the rare combination was in fact higher than its frequent counterpart were the item pairs *what a shame* (.67) and *what an accomplishment* (.74), and *the centre of attention* (.43) and *the ecology of waterways* (.50). As regards to the former pair, it is apparent that changing the rare combination from *what a remarkable player* did not produce the expected result that the more frequent *what a shame* would be easier for the students – in fact the exact opposite happened. This is surprising even considering that Amuzie and Spinner (2013) found that continuous action nouns (*achievement*) were more difficult for their

Korean learners than state nouns (*shame*). It would seem likely that the difference in degree of countability was the more decisive facilitating element for the correct choice of article, compared to that of formulaicity. It is, again, difficult to estimate the effect of removing the adjectival premodification (*remarkable*; which, in theory, should make the item easier) and the change in concreteness (*player* to *achievement*; which, in theory, should make the item more difficult). In addition, one must remember that the positioning of these items was changed in the present test instrument so that the participants would be less likely to become aware of their similarity due to their close vicinity. This may have also contributed to the difference in mean scores, which in this study was 0.07, while only 0.02 in that of Leśniewska (2016).

The items *the centre of attention* and *the ecology of waterways* – which in Leśniewska's results showed no significant difference between their mean scores – were left unchanged because no apparent reason was found relating to the items which might have explained the perceived absence of the facilitating effect of the idiom principle. It is equally challenging to think of why in the present study the mean score for the rarer item was quite higher than in the frequent item. One explanation, which yet again considers the position of the target items inside the test instrument, could nevertheless be considered plausible.

Namely, the item *the ecology of waterways* was the first item in the test, while *the centre of attention* the penultimate one, which may have caused the discrepancy. Assuming that the participants generally started the test from the top, they would have had more time to consider items appearing in the beginning of the test than the later ones due to the limited time in the test situation. Participants not having enough time to reach later test items or becoming fatigued is a common cause for missing data (see for instance Pohl et. al. 2013; Debeer et. al. 2017). In this study there was no specific time limitation for the article filling phase, but participants were reminded to leave some time for the motivation phase as well at approximately halfway through the total time allocated for the two phases. This may have resulted in some students becoming less attentive to later items in the test, including *the centre of attention*. It is impossible to say for certain that this discrepancy resulted from time limitations or participant fatigue, but even such possibility in the present data clearly demonstrates the downsides of the methodology, and naturally jeopardizes the reliability of the results to a certain extent. Issues relating to the methodology are discussed further in chapter 6.

Item pairs that have not yet been touched upon include *twice a day* and *five times a semester*, where there appeared no significant difference in the mean scores in the present study, but

which in Leśniewska's (2016) data showed a slightly significant (p=.04) difference, indicating a small facilitating effect of formulaicity. Judging only by these test items, this usage of the indefinite article would seem rather easy for those students who know the grammar rule, as there is only a small difference in whether the instance appears in a frequent or rare combination. This might, of course, change were different target items used.

Similarly, *speak English* and *learn Kurdish* proved quite easy for the Finnish students, as 96% and 95% of the participants correctly left out all articles in these items respectively. There was a wider gap among the Polish students, of whom the respective 100% and 91% correctly left out all articles, creating a significant (p=.01) difference in the data. Results from these and the previous items highlight the need for further study using larger populations with different educational backgrounds and varied target items in order to confirm whether the effect of formulaicity is indeed significant or not in certain grammatical cases, and if so, among what kinds of learners.

To summarise the analysis of the quantitative data from the article test phase of the study, it seems clear that despite some discrepancies between this replica study and Leśniewska's (2016) original results it can be said with full confidence that there certainly appears to be some sort of psycholinguistic validity for the claim that article processing in some cases is affected by frequency-related mechanisms, which is also supported by previous research on article use and phraseology such as Lenko-Szymańska (2012) and Takahashi (1999) and other research focusing on frequency-driven formulaicity (Ellis 2012). However, the results are also strongly in line with Myles and Cordier's (2017) argument that the dualistic nature of formulaicity must be taken into consideration in formulaic language research. Not all frequent target items in the test, whose role as linguistic clusters was motivated with intuition and corpus-based frequency analysis, exhibited a processing advantage for these participants, and as such, them having a privileged status as some sort of psycholinguistically valid processing unit could not be established. In fact, even for those sequences that did exhibit a processing advantage to be considered processing units, more sophisticated and more psycholinguistically valid testing would have to be conducted, for example by using Myles and Cordier's (2017) suggested methodology.

5.3. Students' Reasons for Their Article Choices

The following chapters will focus on the explanations and reasons that the students wrote when they were asked to motivate their article choices for eight of the target items, which were discussed in chapter 4.3. The analysis is both quantitative and qualitative in nature, the former focusing on the distribution of categorized answers, and the latter on what sort of misconceptions students portrayed in their grammar explanations that were deemed inadequate.

5.3.1. Quantitative Analysis of the Students' Reasons for Their Article Choices

When it comes to these groups of second and third-year Finnish gymnasium students motivating their article choices in the four distinct article rule categories (individuative indefinite reference, individuative definite reference, generic indefinite reference, and a grammatical construct), it would seem that a large proportion of them portrays a lack of metalinguistic knowledge needed for such task. The median for correctly explained items was only two out of the eight items. 20% percent of the students could not correctly explain a single item, and only 8% were able to correctly explain five or more target items. As expected, considering Butler's (2002) similar results, students who received higher grades and were thus placed into the higher proficiency group, on average, portrayed metalinguistic knowledge better than students with lower English grades, very few of whom wrote an adequate explanation to four or more items. Meanwhile over 42% of the higher level students were able to do the same. The distribution of adequately explained reasons can be seen summed in Table 6.

Table 6 Distribution of correctly explained target items.

n adequate explanations	*n* whole group	% whole group	*n* students higher	% students higher	*n* students lower	% students lower
0	21	20,2	3	6,0	18	33,3
1	19	18,3	8	16,0	11	20,4
2	24	23,1	10	20,0	14	25,9
3	15	14,4	8	16,0	7	13,0
4	11	10,6	8	16,0	3	5,6
5	6	5,8	6	12,0	0	0
6	3	2,9	2	4,0	1	1,9
7	3	2,9	3	6,0	0	0
8	2	1,9	2	4,0	0	0
mean	*2,3*		*3,2*		*1,4*	

Students in this study were considerably more likely to achieve higher test score if they were also able to acceptably portray their metalinguistic knowledge in the reasoning phase. Indeed, the number of acceptable answers a student provided correlated considerably positively with overall test scores (.626, p<.001) and with test scores from all the high frequency (.547, p<.001) and all the low frequency items (.633, p<.001). This relation is also illustrated in Figure 1.

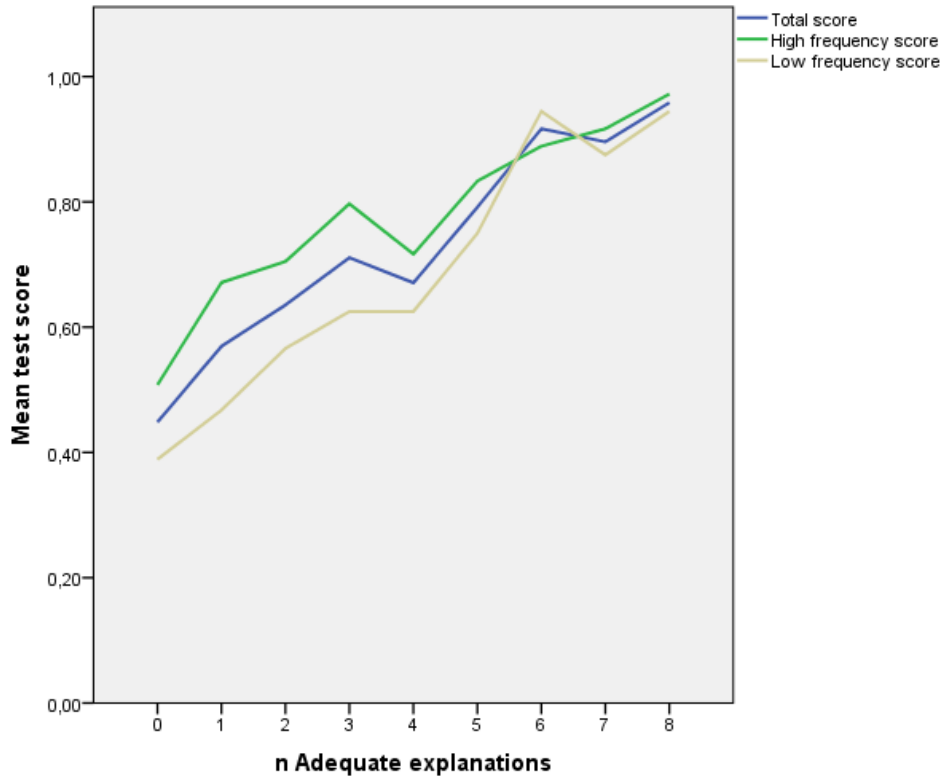

Figure 1 Correctly reasoned article choices in relation to test scores

It is interesting to note that whereas for the lower level students, the respective correlation coefficients for high and low frequency items were very similar, .528, $p<.001$ and .491, $p<.001$ respectively, in the case of the higher level students there was a much weaker correlation in relation to high frequency (.318, $p<.05$) than to low frequency sequences (.495, $p<.001$). In other words, students with higher English grades did not "benefit" as much from their metalinguistic knowledge in the processing of high frequency items, whereas they did so when it came to low frequency items. One possible explanation for this is that they instead benefited from knowledge of the formulaicity of the high frequency sequences and their intuitive judgement, thus mitigating the need for explicit grammar knowledge.

While the above results could be taken as evidence in support of explicit grammar instruction, it must be reminded that correlation does not imply causation, and that the role of explicit and implicit metalinguistic knowledge and language performance has long been an issue of debate in language acquisition studies. The exact nature of these two types of knowledge and how they interact is not clear (Ellis 2009: 20-23; White and Ranta 2002: 261-266), but Ellis (2009) proposes that they be viewed as dichotomous based on neurological evidence. A discussion of the implications of the above correlation between metalinguistic knowledge and language performance unfortunately falls out of the scope of this paper. For one, it cannot answer the question whether these students actually were *utilizing* their grammatical knowledge in the test phase of the study, even if they were able to *portray* it in the reasoning phase. The answers from the reasoning phase are, nevertheless, interesting per se, because they demonstrate the differences in how the frequent and infrequent items were motivated.

Before inspecting the four target item pairs individually, the following table displays the overall distribution of how student responses were categorized. Responses are listed according to targetlike and nontargetlike use and the reader is referred to section 4.6. for a description of the response categories. The percentage shows the proportion in relation to explanations including targetlike and nontargetlike article use.

Table 7 Article choices and response categories – the whole group, higher and lower proficiency students

Response	All students	Higher proficiency	Lower proficiency
Targetlike use			
Adequate	236 (28%)	162 (41%)	74 (17%)
Inadequate	120 (15%)	63 (16%)	56 (13%)
Unattempted	51 (6%)	23 (6%)	29 (6.5%)
Intuitive	99 (12%)	45 (11%)	54 (13%)
Phrasal	54 (6,5%)	30 (7.5%)	24 (5.5%)
Nontargetlike use			
Adequate	9 (1%)	5 (1%)	4 (1%)
Inadequate	46 (6%)	12 (3%)	34 (8%)
Unattempted	118 (14%)	30 (7.5%)	88 (20%)
Intuitive	80 (9.5%)	21 (5%)	59 (14%)
Phrasal	10 (1%)	4 (1%)	6 (1%)
Unintelligible	9 (1%)	5 (1%)	4 (1%)

Generally speaking, the relative number of accepted explanations was the highest among all other categories, but if the numbers from all non-acceptable response categories are summed, they make up the majority (71%) of all answers. This is in stark contrast to the Japanese university students in Butler's (2002) study who were better able to identify the reasons for their article choices, which is to be expected due to them studying at a more advanced level and perhaps them having received more explicit instruction. The total amount of unattempted answers was 20% of both targetlike and nontargetlike usages, which on the one hand can be taken to reflect the participants' lack of metalinguistic competence, or on the other hand, perhaps their hesitation and unwillingness to attempt to portray their knowledge.

Intuition was a common strategy used by the students, and while the lower proficiency group seems to have used intuition to motivate their successful article choices slightly more than students with higher English grades (13% and 11% of the time respectively), their intuitive judgement also failed them more often (14% of the time compared to 5% of the time). In both proficiency groups there was a similar number of students who motivated their choices with the phrase-like quality of the target item, and intriguingly, there were even cases where some students motivated their incorrect article use with such responses. Responses that referred to unknown vocabulary or structure accounted for only 1% of all responses, but they were still important to take into account, as there might have been more students who, instead of writing a response stating that they did not understand the target item, left the item blank.

The following discussion of all four item pairs will display the frequencies and types of responses in each of the eight target items. The objective was to find out to what extent the participants were consistent in their grammar explanations in relation to items with differing formulaicity yet with same underlying grammar. To begin with, Table 8 shows the response frequencies for *have kids* and *eat carbohydrates*, which were generally speaking very easy for the participants in that almost all of them correctly did not use any article with these phrases. Still, a small gap of 0.06 was present between the mean scores for these items.

Inspecting the student responses on a general level, there seemed to be no noticeable differences in the way students motivated these two items, other than that students were more successful in giving an acceptable explanation for *have kids* than for *eat carbohydrates*, and that students relied on intuition slightly more with the former. One student who mistakenly wrote *have a kids* motivated this choice by saying he or she remembered such phrase as a whole, which might indicate some sensitivity to the 3-gram *have a(n)* X (see discussion in section 4.3.). It is difficult to say whether the small gap of 0.06 in mean scores resulted from formulaicity effect of *have kids* or whether there was some other factor. The analysis of inadequately explained answers in 5.2.2. hopes to shed light on this matter.

Table 8 Student responses for *have kids* and *eat carbohydrates*

Response	*have kids*	*eat carbohydrates*
Targetlike use		
Adequate	25 (24%)	19 (18%)
Inadequate	38 (37%)	44 (42%)
Unattempted	15 (14%)	12 (12%)
Intuitive	21 (20%)	17 (16%)
Phrasal	-	-
Nontargetlike use		
Adequate	-	-
Inadequate	1 (1%)	6 (6%)
Unattempted	0	1 (1%)
Intuitive	3 (3%)	5 (5%)
Phrasal	1 (1%)	-

In the items *the sooner the better* and *the smaller the pot the more critical the problem* (summarized in Table 9) the student responses displayed much more variance than in the previous pair of items. What is noticeable is the large number of responses referring to phrasal quality of the item *the sooner the better*. Over one third of the students motivated their article choices in this item by stating that it was a phrase or by referring to its holisticity. This is not

surprising at all, given the frequency of *the sooner the better* (0.29 and 0.28 occurrences per million words in the COCA and BNC respectively) and the likelihood for it to be given as an example when instructing the use of the *the – the* construct with adjectives. There were even some mentions of phrasality for the infrequent item *the smaller the pot the more critical the problem*, but these were cases where the student had written "same as above" in their response, referring to *the sooner the better*. These students had perhaps meant to refer to the *the – the* construct as the phrasal element, and not *the sooner the better* per se, although this is difficult to ascertain.

Table 9 Responses for *the sooner the better* and *the smaller the pot the more critical the problem*

Response	*the sooner the better*	*the smaller the pot the more critical the problem*
Targetlike use		
Adequate	21 (20%)	15 (14%)
Inadequate	1 (1%)	1 (1%)
Unattempted	2 (2%)	5 (5%)
Intuitive	12 (11%)	10 (10%)
Phrasal	34 (32%)	3 (3%)
Nontargetlike use		
Adequate	2 (2%)	2 (2%)
Inadequate	6 (6%)	7 (7%)
Unattempted	10 (10%)	33 (31%)
Intuitive	11 (11%)	22 (21%)
Phrasal	5 (5%)	2 (2%)
Unintelligible	-	4 (4%)

There were four responses referring to the unintelligibility of the item *the smaller the pot the more critical the problem* as a source for not having chosen any articles in the test phase. This seemed to be due to the complexity of the item and not the vocabulary, which was quite basic. These four students were still able to insert the correct articles for *the sooner the better*, but only one of them gave an adequate grammar explanation in his or her response, and the three others referred to either intuition or phrasal quality. Out of the other 69 students who accurately inserted the articles in the common phrase *the sooner the better*, 32 could and 37 could not do the same for the rare combination. As it turns out, the former 32 students also recognized the grammar for either one of these two items much more often than the latter 37 students, many of whom relied on intuition or memory when dealing with the items. 17 of the former 32 and 7 of the latter 37 respondents recognized the grammar.

Overall, as regard to this item pair, the participants seemed to be greatly assisted by the formulaicity of *the sooner the better*, which is evident from the number of responses that referred to it as a phrase. But relying solely on this knowledge was not so helpful when it came to the item *the smaller the pot the more critical the problem*, since many who did so in the frequent item failed in the rare item.

It is also interesting that out of the 39 participants who relied on phrasality of the item *the sooner the better*, five still failed to insert the correct articles. Four of them wrote *sooner the better* and one *the sooner better* instead. These seem like instances where the participant's memory trace of the phrase included an unstressed *the* in one of the positions, which is understandable given that articles are rarely stressed in speech, which greatly contributes to their difficulty (Master 2002: 332).

Frequencies for the different kinds of responses for *the day I die* and *the food I brought*, shown in Table 10, were also quite similar besides the fact that sometimes the participants regarded *the day I die* as a holistic unit (13% of the time) and that there were more unaccepted and unattempted explanations for *the food I brought*. A statistically significant gap of 0.21 between the mean scores for these items was found in the previous section, and it was expected that just like with the previous items discussed, this could have mainly resulted from students relying on intuitive judgement and long-term memory in the case of *the day I die*, but failing to consider the grammar in *the food I brought*. However, in this case the examination showed that there was no considerably higher reliance on intuition or memory in case of those students who were accurate in *the day I die* but failed in *the food I brought*. As with all the other item pairs, reasons for the students' failure to give an acceptable grammar explanation will be discussed in the next section, which might reveal something that resulted in more inadequate responses for the item *the food I brought*.

Table 10 Responses for *the day I die* and *the food I brought*

Response	*the day I die*	*the food I brought*
Targetlike use		
Adequate	67 (64%)	55 (53%)
Inadequate	4 (4%)	8 (7%)
Unattempted	1 (1%)	3 (3%)
Intuitive	12 (12%)	10 (10%)
Phrasal	14 (13%)	-
Nontargetlike use		
Adequate	-	1 (1%)
Inadequate	1 (1%)	11 (11%)
Unattempted	4 (4%)	9 (8%)
Intuitive	1 (1%)	7 (7%)
Phrasal	-	-

Lastly, a comparison between the responses for *a friend of mine* and *an acquaintance of mine* revealed that the students' motivations were drastically different for these two items (Table 11). The former item was significantly easier for the participants to process, resulting in a gap of 0.34 between the mean scores when only 27% of all participants correctly inserted *an* (or *a*) in the phrase *an acquaintance of mine* compared to the accuracy of 61% for *a friend of mine*. The majority of incorrect answers resulted from *zero* answers. This was despite the fact that the items appeared very similar in syntax (*a(n)* X *of mine*).

Table 11 Responses for *a friend of mine* and *an acquaintance of mine*

Response	*a friend of mine*	*an acquaintance of mine*
Targetlike use		
Adequate	23 (22%)	11 (11%)
Inadequate	18 (17%)	5 (5%)
Unattempted	7 (7%)	7 (7%)
Intuitive	13 (13%)	4 (4%)
Phrasal	3 (3%)	-
Nontargetlike use		
Adequate	2 (2%)	2 (2%)
Inadequate	7 (7%)	7 (7%)
Unattempted	17 (16%)	44 (42%)
Intuitive	12 (11%)	19 (18%)
Phrasal	2 (2%)	-
Unintelligible	-	5 (5%)

Almost half (42%) of the respondents did not attempt any grammatical explanation for *an acquaintance of mine* and the number of explanations that demonstrated the participants' understanding of the grammar was half of that of *a friend of mine*. Five students wrote that they did not understand the word *acquaintance*, and the large number of unattempted responses indicates that there might have been more unstated problems in comprehension. It is difficult to think of another explanation, given that in Herranen's (1977) and Lehtonen's (2015) papers the indefinite reference was mastered relatively well by Finnish students.

In summary, this section examined the overall picture of how the participants approached the eight target items when prompted to explain their own article choices for these items. Students had to explain their article use in generic, indefinite and definite references and in the comparative correlative construction *the – the* in their own words, whether they had or had not inserted any articles in the phrase. The definite reference was clearly the easiest to explain in grammatical terms compared to the other items. Comparison between responses for the frequent and infrequent items showed that it was considerably harder for participants to provide acceptable explanations for the infrequent items, and that they instead left these items more often either completely blank or without stating any motivation other than "I guessed" or "I don't know". Prior to the analysis, it was hypothesized that the opposite could have taken place; that the students would tend to motivate their choices for frequent items with intuition, whereas the infrequent items would have required more explicit thought. But it must be remembered that the test instrument instructed students to first attempt explaining the rule, which may have affected the results. Still, that a large number of responses referred to formulaicity especially in the items *the sooner the better* and *the day I die* did show how these items were recalled as holistic processing units, although sometimes incorrectly, and that on average, students' intuition was more accurate in the case of high frequency items. Infrequent items were neither regarded as holistic units nor did the students' intuition help as much in their processing.

In some of the items, unattempted explanations accounted for a large part of all answers, especially in those cases of nontargetlike article use. It is impossible to know how many students realized during the reasoning phase that their article choice in the test phase was incorrect and therefore chose not to write any explanation for what they knew was incorrect. Or perhaps the items were simply too demanding in vocabulary and syntax that the students were unable to give any reason. The latter seems unlikely, because the vocabulary was, despite the challenges discussed above, quite basic and something gymnasium students were expected

to be familiar with. In any case, it would have been wise to encourage the participants to attempt an explanation regardless of if they thought their choice in the test phase was correct or not.

It is unfortunate that the present analysis had to be limited to only the above eight target phrases, because the resulting picture is undeniably very incomplete. It would have been interesting to compare student responses for item pairs such as *get a job* and *get a wage increase* and see whether an equally diverse set of responses would surface as with *a friend of mine* and *an acquaintance of mine*, despite syntactic similarities.

5.3.2. Qualitative Analysis of the Students' Unacceptable Reasons for Their Article Choices

This section assumes a qualitative approach in the analysis of the students' written responses in the reasoning phase of the study. More precisely, the focus is on answers that were categorized as inadequate, and the question of interest is: what sort of misconceptions about the English article system are reflected in the students' responses? The purpose is exploratory and meant to complement the data already presented in this paper; while the metalinguistic competence of these young learners of English was not the main concern in this study, such analysis can nevertheless shed more light on the processing of the particular test items students were expected to analyse grammatically. In addition, the analysis will highlight some further shortcomings in the methodology and test instrument used.

Because other categories of responses (explanations deemed adequate, answers referring to reliance on intuition, answers referring to the formulaicity of the target item, answers left blank or without a reasoning, and those stating that the target item was not understood) were more self-explanatory in nature, an in-depth analysis of those categories was considered redundant, although it certainly would be interesting to see *how* students express their metalinguistic knowledge with metalingual means as, for example, Hu (2011) examined. All the examples have been translated (as literally as possible) from Finnish into English by the researcher.

First, the item pairs *eat carbohydrates* and *have kids*, which had the highest amount of inadequately explained responses will be discussed. While going through the data it immediately became obvious that the majority of student replies that failed to meet the criteria of an adequate explanation in both these items consisted of students only mentioning the plurality of either *carbohydrates* or *kids*, thus failing to mention their indefiniteness. All these

answers are, of course, not wrong per se; *carbohydrates* and *kids* do refer to plural entities, and in this case no harm came to these students' test score because not inserting an article was correct. However, as explained in section 4.6., both nouns can also be referred to as definite entities even in plural form, so explicit mention of their indefiniteness was considered a requirement. Examples (1) and (2) show typical replies for these items.

(1) "'Carbohydrates' is a plural, no article needed."

(2) "No article with a plural."

Eat carbohydrates and *have kids* had 53 and 39 inadequate explanations in total respectively, and these sorts of instances were found 30 times in each item often, but not always, written by the same participant.

A literal interpretation of responses equivalent to example (2) would indicate that the student writing so would mistakenly never insert an article before a plural noun. However, it is difficult to know whether this is the actual case or whether it simply was a failure to consider the possibility of definite reference. Had there been other target items where the students needed to produce an explanation for a definite plural noun phrase, these students would have perhaps altered their reasoning. There was also one participant who mistakenly identified *food* in *the food I brought* as plural and wrote a similar explanation as (2) for this item as well. This is also one shortcoming of the present methodology; the responses student wrote were very brief, just as instructed, and interaction with the researcher was quite limited, and thus there was no possibility for requesting clarification.

Other less frequent hypotheses for the item *eat carbohydrates* served as a reminder how carefully a test instrument such as in the present study must be considered in order to avoid factors undermining the reliability of the results. The issue of target item pairs differing in boundedness and countability was already discussed in section 4.4., and the same issue is also present here. Namely, *carbohydrates* can be used as a both countable and non-countable noun to refer to the class of carbohydrates or its substance, while *kids* is a concrete countable noun. This resulted in the item *eat carbohydrates* receiving a few responses like in example (3), referring to its substance-like quality. Results from the test phase did not indicate that this was a large-scale problem in this study, as the mean scores for these items were very similar, but it is crucial that future research takes such, even small factors into consideration.

(3) "'Carbohydrates' is a substance word, no article."

The examples above were cases where the participant had correctly not used any article before the noun phrases, but there was a small number of instances–where the participant had inserted an article and attempted to explain their reasoning. Some had mistakenly identified *carbohydrates* or *kids* as definite, stating, for example, that they were "specific 'parts' of food/nutrition" or that "the kids had already been mentioned". One participant who had inserted the indefinite *a* for both these items had explained, correctly, that the target items in question were nouns, which would indicate his or her failure to consider the plurality of the items. There were very few other responses of this kind, and these respondents' school grades and test scores were below average, suggesting that English language in general and particularly the article system was difficult to them.

Secondly, the item pair *the sooner the better* and *the smaller the pot the more critical the problem* required knowledge of the grammatical function of *the – the* construction. These items had the most variance in the way the participants had filled in the articles in the test phase, sometimes leaving out all articles, sometimes only one, or sometimes using the indefinite article inappropriately inside the phrases. Two types of problems arise from the students' responses in these items: first, students exhibited a lack of knowledge of the function of this construction, explaining why they had not inserted any articles. Explanations like (4) were found a couple of times in *the sooner the better* and (5) in both items.

(4) "There are no nouns."

(5) "*The* is needed before a superlative, but not with a comparative".

The second type of mistake in the case of *the smaller the pot the more critical the problem*, some students were confused by the more demanding syntax, attempting to explain how the articles work with the noun phrases *pot* and *problem* blended inside the phrase instead of the *the – the* construct. This, again, highlights the fact that formulaicity most likely was not the only factor that resulted in the large gap between mean scores for these items, because an item with a more complex syntactic context can inherently be more difficult for learners. Thus, when instructing the use of the *the – the* construct, attention should be paid on more complex sentences such as the above item.

Thirdly, an examination of the responses for the items *the day I die* and *the food I brought* revealed yet again how the difference in countability might have affected the item scores. In the latter item's case, out of the 19 inadequate explanations 6 attempted to explain that *food*

was non-countable or a substance word, and thus it could not take an article. All such answers were very similar to example (6).

(6) "Food cannot be counted -> no article."

It is, again, difficult to interpret from responses such as (6) whether the respondents immediately rejected the possibility of all articles after detecting the referent as non-countable, or whether they meant that no indefinite article could be used. These students had probably been instructed that mass nouns used in non-countable sense never take an indefinite article, and were either overgeneralizing the rule or failing to consider the item's definiteness. Naturally, no such responses were written in *the day I die*, as *day* is a countable noun.

Rather, the rest of the student explanations for both *the day I die* and *the food I brought*, consisted mostly of idiosyncratic hypotheses (examples 7, 8) or references to word-article collocational rules (examples 9, 10), similar to what Butler (2002: 468) also found among the Japanese students' responses. Butler herself suggests that the latter type of hypotheses might be the result of English articles occurring frequently in such bundles containing prepositions, just as Lenko-Szymańska's (2012) corpus analysis revealed. It would appear that some of the Finnish participants were also sensitive to these kinds of lexical bundles.

(7) "That particular day is being stressed here."

(8) "*The* comes before an expression of time."

(9) "*The* comes automatically after 'until'?"

(10) "*Any of* requires the *the* article."

Finally, the items *a friend of mine* and *an acquaintance of mine*, the latter of which was clearly more difficult for the participants to give a grammatical explanation for, as was seen in the previous section. There were twice as many unattempted explanations and inadequate explanations for the latter item. This can be considered surprising, though, as the grammar rule and syntactic environment for both items were similar. Collocational hypotheses – which, coming from different students could be quite contradictory to each other – surface here too, as examples (11), (12) and (13) illustrate.

(11) "*of mine* possessive construction requires the definite article THE."

(12) "After the word *by* comes no article."

(13) "Word *by* is followed by *a* or *an*?"

Another major problem in the students' responses was the inappropriate use of the word "tietty", which in English translates to "specific" or "certain" (e.g. "tietty ystävä", "a certain/specific friend"), and was interpreted to mean "definite" in this context. Eleven students in total attempted to explain that either friend or acquaintance or both in their respective test sentences were *tietty*, but still all these students but one inserted the indefinite article, which is used when the referent is the opposite of *tietty*, that is, unknown or uncertain. It is surprising that so many students, with above average English grades too, used the term mistakenly, and it might reflect their confusion of all the metalingual terms related to noun phrases and reference.

It is also difficult to confirm why *an acquaintance of mine* had so many fewer attempted explanations than *an acquaintance of mine*. The syntactic environment was almost identical (*An acquaintance of mine* starting a sentence and *a friend of mine* acting as an agent preceded by *by*, resulting in hypotheses such as 12, 13) so it was expected that they would have less variance. However, in the previous section it was reported that *an acquaintance of mine* received five responses from students who did not know the word *acquaintance* and expressed that this was the reason they failed to insert the correct article in the test phase. It is possible that this was the case for some other students as well, but instead of noting this in their response they chose to leave it blank. In hindsight, adding *acquaintance* to the list of words whose Finnish translation was provided in the test phase (see section 4.3.) might have been a sound decision. There seems to be nothing else indicating the reason for the apparent difficulty of the item *an acquaintance of mine*, so the conclusion is made that the divergence in mean scores resulted, on the one hand, from formulaicity effect, and, on the other hand, from unintelligibility of the word *acquaintance*. The exact extent of these factors remains unknown.

This data was gathered and analysed for the purpose of gaining a more in-depth picture of the students' metalinguistic knowledge when prompted to provide a grammatical explanation for some of their article choices. A brief analysis of those explanations that did not meet the criteria for an adequate grammatical explanation showed that the students either failed to elaborate their answer enough for it to be considered acceptable or portrayed various kind of misconceptions about the article system. Sometimes students portrayed sensitivity to certain word combinations that frequently occur with articles, which can be said to both hinder and facilitate article processing. The analysis subsequently revealed further problems in the test methodology mainly concerning additional cases of the target item pair differing in boundedness or syntactic

context. Leśniewska (2016) also contemplated whether such factors affected her test results among Polish students, and this seems likely. At least in the present study it is evident, given that the participants gave different and inconsistent explanations for the target item pairs, sometimes referring to other sources of difficulty in the case of infrequent items.

6. Conclusion

Four research questions were introduced in section 4.1. of this thesis, and the analysis in chapter 5 was carried out in accordance to them. First, concerning the first two research questions, this paper set out to investigate the claim about article processing being facilitated by frequency-related mechanisms and the idiom principle, which was first presented by Leśniewska (2016) in her study on Polish learners of English and their article use in situations where articles occurred in either frequent or rare sequences of lexis but which were syntactically similar. Just as with the Polish university students, the Finnish gymnasium students' article use in this paper was significantly more accurate in frequent combinations compared to infrequent sequences. Divergent results were also bound to arise due to different study backgrounds and the proficiency gap between the Finnish and Polish students. Namely, while Leśniewska (2016) considered nine out of twelve target pairs in her results to exhibit the facilitating effect of formulaicity based on the statistical significance in mean score differences, the same analysis resulted in only five of the twelve items and their mean differences being statistically significant in the present study. But even though differences in the rest of the combination pairs was not statistically significant, in ten pairs the mean score for the frequent item was, nevertheless, higher.

Secondly, the aspect of metalinguistic knowledge was reflected in the third and fourth research questions, which were concerned with the metalinguistic knowledge that the students portrayed, and what differences and misconceptions their responses showed between frequent and rare target items. As for the metalinguistic knowledge that the Finnish students portrayed in their explanations, it was evident that many lacked the explicit knowledge needed to verbalize an acceptable grammatical explanation regarding their article use. Those who could express their knowledge in an adequate manner tended to perform significantly better in the article filling test, but despite this correlation, a causational relationship could not be established (and was not even attempted) due to limitations in this study. Motivations for the four target item pairs did, however, reveal some interesting facts about how the participants approached these items in the reasoning phase and how this might have affected their processing of the item sequences in the test phase. For one, some items were clearly regarded as holistic units by many of the participants, which may have helped them choose the correct article for those sequences in the test phase even if they could not state the grammar for these items or their rare counterparts.

This point of view does argue for the fact that explicit instruction is needed in order to prepare students for those instances of article use that are novel and unfamiliar.

The above results align perfectly with other research on formulaicity that has showed a processing advantage for formulaic language (such as Ellis 2012, Schmitt and Underwood 2004, Tremblay and Baayen 2010) and even research that has touched upon article use in conventionalized situations (Leńko-Szymańska 2012, Takahashi 1997). A trend clearly seems to exist, and although it is still far too early to make broader generalizations, it is proposed here that other populations of non-native English learners would most likely exhibit a similar sensitivity to formulaicity in their article processing too, and therefore further research is crucial. There is more than enough reason to assume that not all instances of article use during discourse are generated online, as the analytic view of language purports, but rather processed according to implicit knowledge and the idiom principle in a similar manner as other stretches of language that display characteristics of processing units.

An important theoretical question arises hence: why limit this discussion to only English articles? Leśniewska (2016) and the present paper suggest a view of formulaicity integrating lexis with a very small section of grammar (articles), but if formulaicity of language is as pervasive as for example Conklin and Schmitt (2012) maintain, should not this view be made to incorporate all grammatical aspects of a language? Take, for example, English prepositions or the Finnish case system, which are known for their intricacies. Other possible questions for future research could be whether formulaicity affects the use of prepositions in a facilitating manner among native and second language speakers of English, or to which extent the fluent use of the Finnish noun case system relies on processing units. The possibilities seem endless should the phraseological perspective be adopted in a wider spectrum of linguistic research.

As illustrated many times in the present paper, future research must make certain that the construct of formulaicity is valid and that if item pairs of frequent and rare (or formulaic and nonformulaic) combinations are used, such as in the present paper, all care must be taken to ensure that no external factors of difficulty are present, such as unknown vocabulary, premodification or postmodification, or difference in the abstractness of noun phrases. The student explanations illustrated that in the present study this was the case in some target items, which does reduce the reliability of the results. Future studies must also take advantage of other more psycholinguistically valid test instruments. Eye-tracking methods, oral production or electrophysical ERP data would all illustrate the psycholinguistic validity of formulaic language

better than a written test. If a written test similar to the one in the present paper and Leśniewska (2016) is used, at the very least, the positioning of the test items inside the test instrument should be considered in order to avoid some items having inadvertent saliency and to mitigate the effect of participant fatigue on the results. Not to mention that the time constraints in the test setting should be eliminated, or at least controlled.

The view argued in this paper does hold some pedagogical implications for article (and perhaps other grammar) instruction. As mentioned above, instructors should try to ensure that their students are equipped for instances of article use that fall outside of conventional use, whatever the methods of instruction they choose. But on the other hand, formulaicity can also help automatize article use in rarer cases, such as proper nouns and geographical areas. The most archetypical examples of even peripheral uses of articles could be highlighted. Furthermore, it was noted that phonological factors can have an effect on the difficulty on article acquisition, because they rarely occur in stressed positions. This could lead into incorrect memory traces of even formulaic sequences, such as responses like *sooner the better* shows. Thus, instruction might benefit from clear articulation of articles in classroom discourse.

The phraseological perspective has also some very important implications for language testing. An integral part of language testing and test instrument design is the test construct, which refers to the knowledge and skills of the test subject that are intended to be measured with the test, and how these skills are reflected in language performance (McNamara 2000: 13). Say that an instructor wanted to test his or her subjects on article use and designed a test instrument to measure how accurately they are able to fill in articles in an articleless text. In the test, the instructor would certainly want to include target items that the subjects would need to process according to the open-choice principle, and not the idiom principle. Formulaicity in this case would introduce variance that is construct irrelevant (McNamara 2000: 53), because the goal is to measure the subjects' article processing ability and not their long-term memory. This does make test design an even more challenging task than it already is, and besides, a test designer following this suggestion would in all practicality be limited to his or her intuition in determining what in the test constitutes as formulaic and what does not, for he or she cannot know beforehand what psycholinguistic units the subjects possess. But at the very least, the designer could avoid using target items that have been discussed during recent instruction, for example, in classroom texts. Such sequences are perhaps more likely stored in the subjects' memory as holistic units.

Bibliography

Amuzie, G., L. and Spinner, P. (2013). Korean EFL learners' indefinite article use with four types of Abstract Nouns. *Applied Linguistics* 34 (4), 415-434. doi: 10.1093/applin/ams065.

Akakura, M. (2012). Evaluating the effectiveness of explicit instruction on implicit and explicit L2 knowledge. *Language Teaching Research* 16 (1), 1-29. doi: 10.1177/1362168811423339.

Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Longman.

Conklin, K. and Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics* 32, 45-61. doi: 10.1017/S0267190512000074.

Cowie, A. P. (1998). *Phraseology: Theory, analysis and applications*. Oxford: Clarendon Press.

Davies, M. (2008-) *The corpus of contemporary American English: 450 million words, 1990-present*. Retrieved from http://corpus.byu.edu/coca/.

Debeer, D., Jannsen, R. and De Boeck, P. (2017). Modeling skipped and not-reached items using IRTrees. *Journal of Educational Measurement* 54 (3), 333-363.

DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning* 55, 1–25. https://doi.org/10.1111/j.0023-8333.2005.00294.x.

Downing, A. and Locke, P. (2006). *English grammar: A university course.* 2nd ed. New York: Routledge.

Ekiert, M. (2004). Acquisition of the English article system by speakers of Polish in ESL and EFL settings. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics* 4 (1). doi: 10.7916/D81C28DK.

Elder, C. (2009). Validitating a test of metalinguistic knowledge. In R., Ellis, S., Loewen, C. Elder, R. Erlam, J. Philip and H. Reinders (eds.) *Implicit and explicit knowledge in second language learning, testing and teaching*, 113-138.

Ellis, N. (2012). Formulaic language and second language acquisition: Zipf and the phrasal teddy bear. *Annual Review of Applied Linguistics* 32, 17-44. doi: 10.1017/S0267190512000025.

Ellis, N. and Simpson-Vlach, R. (2009). Formulaic language in native speakers: Triangulating psycholinguistics, corpus linguistics, and education. *Corpus Linguistics and Linguistic Theory* 5, 61-78. doi: 10.1515/CLLT.2009.003.

Ellis, N., Frey, E. and Jalkanen, I. (2009). The psycholinguistic reality of collocation and semantic prosody (1). In U. Römer and R. Schulze (eds.) *Exploring the lexis-grammar interface* 89-114. Amsterdam: John Benjamins.

Ellis, N., Simpson-Vlach, R., Maynard, C. (2008). Formulaic language in native and second language speakers: Psyscholinguistic, corpus linguistic, and TESOL. *TESOL Quarterly* 42 (3), 375-396. https://www.jstor.org/stable/40264474.

Elder, R. (2009). Implicit and explicit learning, knowledge and instruction. In R., Ellis, S., Loewen, C. Elder, R. Erlam, J. Philip and H. Reinders (eds.) *Implicit and explicit knowledge in second language learning, testing and teaching*, 113-138.

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

Hakulinen, A., Vilkuna M., Korhonen, R., Koivisto, V., Heinonen, T. R. and Alho, I. (2004). *Iso suomen kielioppi*. 2nd ed. Hämeenlinna: Karisto Oy.

Harb, M. A. (2014). A closer look at the English article system: Internal and external sources of difficulty revisited. *International Journal of Linguistics* 6 (4), 87-101. doi: 10.5296/ijl.v6i4.5998.

Herranen, T. (1977) *Errors made by Finnish university students in the use of the English article system*. Unpublished Pro Gradu Thesis. University of Jyväskylä, Department of English.

Juvonen, P. (2000). *Grammaticalizing the definite article. A study of definite adnominal determiners in a genre of spoken Finnish*. Doctoral Thesis. Stockholm: Deptartment of linguistics.

Lamb, S. M. (1998). *Pathways of the brain: The neurocognitive basis of language*. Amsterdam: John Benjamins.

Langacker, R. W. (1991) *Foundations of cognitive grammar. Vol. II: Descriptive application*. Stanford CA: Stanford University Press.

Larjavaara, M. (2001). Määräinen artikkeli – suomessa? [online]. *Kielikello* 4/2001. https://www.kielikello.fi/-/maarainen-artikkeli-suomessa. (1.2.2018).

Lehtonen, P. (2015). *Referens och artikelbruk på olika cefr-nivåer hos finskspråkiga universitetsstudenter - en jämförelse mellan svenska och engelska.* Unpublished Pro Gradu thesis. University of Jyväskylä, Department of Languages. http://urn.fi/URN:NBN:fi:jyu-201506102272.

Leńko-Szymańska, A. (2012). *The role of conventionalized language in the acquisition and use of articles by Polish EFL learners*. Warsaw University.

Leśniewska, J. (2017). New perspectives on the use of articles in L2 English. In L. Torres-Zúñiga and T. H. Schmidt (eds.) *New methodological approaches to foreign language teaching*, 67-82. Cambridge Scholars Publishing.

Leśniewska, J. (2016). The use of articles in L2 English: A phraseological perspective. In *Studies in Second Language Learning and Teaching* 6 (2), 203-224.

Li, H. and Yang, L. (2010). An investigation of English articles' acquisition by Chinese learners of English. *Chinese Journal of Applied Linguistics* 33 (3), 15-31.

Master, P. (1997). The English article system: Acquisition, function, and pedagogy (1997). *System* 25 (2), 215-232.

Master, P. (2002). Information structure and English article pedagogy. *System* 30, 331-348.

Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics* 32, 111-129. doi: 10.1017/S0267190512000128.

Moon, R. (1998). *Fixed expressions and idioms in English: A corpus based approach*. Oxford: Clarendon Press.

Myles, F. and Cordier, C. (2017). Formulaic sequence (FS) cannot be an umbrella term in SLA: Focusing on psycholinguistic FSs and their identification. *Studies in Second Language Acquisition* 39, 3-28. doi: 10.1017/S027226311600036X.

Opetushallinnon tilastopalvelu Vipunen (2016). Kieli- ja muut ainevalinnat. https://vipunen.fi/fi-fi/perus/Sivut/Kieli--ja-muut-ainevalinnat.aspx (20.5.2018).

Pawley, A. and Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards and R. W. Schmidt (eds.), *Language and communication*. New York: Longman, 191−226.

Pohl, S., Gräfe, L. and Rose, N. (2014). Dealing with omitted and not-reached items in competence tests: Evaluating approaches accounting for missing responses in item response theory models. *Educational and Psychological Measurement* 74 (3), 423-452. doi: 10.1177/0013164413504926.

Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman.

Radden, G. and Dirven, R. (2007). *Cognitive English grammar*. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Read, J. and Nation, P. (2004). Measurement of formulaic sequences. In Schmitt, N. (eds). *Formulaic Sequences. Acquisition, processing and use.* Amsterdam: John Benjamins, 23-35.

Richard, A., B. (2004). *Regression Analysis: A Constructive Critique*. Sage Publications. doi: 10.4135/9781483348834

Schmitt, N. and Underwood, G. (2004). Exploring the processing of formulaic sequences through a self-paced reading task. In N., Schmitt (eds.), *Formulaic sequences. Acquisition, processing and use.* Amsterdam: John Benjamins, 173-189.

Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford, UK: Oxford University Press.

Sinclair, J. M. (2004). *Trust the text: Lexis, corpus, discourse*. London: Routledge.

Taanila, A. (2014). Määrällisen aineiston kerääminen. http://myy.haaga-helia.fi/~taaak/t/suunnittelu.pdf (20.1.2018).

Takahashi, T. (1997). Japanese learners' acquisition and use of the English article system. *Edinburgh Working Papers in Applied Linguistics* N8. https://files.eric.ed.gov/fulltext/ED409716.pdf.

The British National Corpus, version 3 (BNC XML Edition). 2007. Distributed by Bodleian Libraries, University of Oxford, on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk/

Trenkic, D. (2007). Variability in second language article production: beyond representational deficit vs. processing constraints debate. *Second Language Research* 23 (3), 289-327.

Tremblay, A. and Baayen, H. (2010). Holistic processing of regular four-word sequences: A behavioral and ERP study of the effects of structure, frequency, and probability on immediate free recall. In D. Wood (eds.), *Perspectives on formulaic language: Acquisition and communication*, 151–173. London, UK: Continuum International.

Vartiainen, M. (1979) *The development of the article system in the English of Finnish schoolchildren*. Unpublished Pro Gradu Thesis. University of Jyväskylä, Department of English.

Walker, R. (2004) *Getting and analysing of qualitative data*. The PREST Training Resources. Commonwealth of Learning. oasis.col.org/bitstream/handle/11599/87/A4.pdf

White, L. (2003). Fossilization in steady state L2 grammars: Persistent problems with inflectional morphology. *Bilingualism: Language and Cognition* 6 (2), 129-141. doi: 10.1017/S1366728903001081.

White, J. and Ranta, L. (2002) Examining the interface between metalinguistic task performance and oral production in a second language. *Language Awareness* 11 (4), 259-290. doi: 10.1080/09658410208667060.

Wood, D. (eds.) (2010). *Perspectives on formulaic language: Acquisition and communication*. London: Continuum.

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics* 21 (4), 463–489.

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

Wray, A. (2008). *Formulaic language: Pushing the boundaries*. Oxford: Oxford University Press.

Wray, A. (2012). What do we (think we) know about formulaic language? An evaluation of the current state of play. *Annual Review of Applied Linguistics* 32, 231-254. doi: 10.1017/S026719051200013X.

Wray, A. (2013). Formulaic language. *Language Teaching* 46 (3), 316-334. doi: 10.1017/S0261444813000013.

Zdorenko, T., Paradis, J. (2011). Articles in child L2 English: When L1 and L2 acquisition meet at the interface. *First Language* 32 (1-2), 38-62. doi: 10.1177/0142723710396797.

Appendix

A: *Instrument for the test phase*

OHJE: Alla olevassa tekstissä ei ole artikkeleita. Lisää artikkelit (a, an, the)**. Mikäli artikkelia ei mielestäsi tarvita, älä merkitse mitään.**

Esimerkki: I always play the piano on Saturdays.

1. Motorboats harm ecology of waterways, unless their use is kept at low level.

2. Glucose, or blood sugar, is produced in our bodies when we eat carbohydrates.

3. We meet regularly, five times semester, at departmental meeting.

4. Time matters. Please try to send it in as soon as possible - sooner better.

5. I want to choose foreign language that few people want to study. Maybe I'll learn Kurdish.

6. Plants in pots and containers require more water than you actually might think, smaller pot more critical problem. By midsummer, herbs and vegetables in containers may need water twice day.

7. You should give him spoonful of this syrup every three hours.

8. I'll remember you until day I die.

9. I see that you haven't eaten any of food I brought you two days ago. Can I make you cup of tea?

10. Old leftist political parties are re-emerging to demand that government again expand its role in economy to help poor, even at price of discouraging foreign investors.

11. Immediately after graduation I need to get job. It doesn't necessarily have to be in my field, and I'm prepared to move anywhere where I can find work. Acquaintance of mine was recently offered position in Berlin and he moved there without moment's hesitation.

12. I was lucky ball didn't hit me in face. What shame it got lost, though.

13. New version of insurance policy makes number of alternatives open to insured.

14. Do you speak English?

15. I was recently asked about my hopes for future by friend of mine. What I know is that I'd like to have kids. And I'd like to get wage increase some time soon.

16. Best player in their team scored fifty home runs during season. What accomplishment it was.

17. Every member of Royal Family enjoys star status; they are used to being centre of attention and there is strong unstated rivalry between them.

18. He was cut in hand in same fight, according to testimony.

B: *Instrument for the reasoning phase*

Ohje: Perustele lyhyesti seuraavat artikkelivalintasi omin sanoin. Täytä myös sivun alalaitaan kolmen viimeisimmän englannin kurssisi numerot, ja mikäli olet oleskellut ulkomailla.

Perustellessasi muista, että ”en tiedä, kielikorva sanoi”, ”arvasin” yms. perustelut hyväksytään, mutta yritä ensin miettiä kieliopillista selitystä.

a. Katso artikkelivalintaasi kohdassa **2**: **… whenever we eat ___ carbohydrates.**
Miten perustelet artikkelivalintasi alleviivatussa kohdassa?

---

b. Katso artikkelivalintojasi kohdassa **4**: **… __ sooner __ better.**
Miten perustelet artikkelivalintasi alleviivatuissa kohdissa?

---

c. Katso artikkelivalintojasi kohdassa **6**: **… ___ smaller pot, ___ more critical problem.**
Miten perustelet artikkelivalintasi alleviivatuissa kohdissa?

---

d. Katso artikkelivalintaasi kohdassa **8**: **I’ll remember you until ___ day I die.**
Miten perustelet artikkelivalintasi alleviivatussa kohdassa?

---

e. Katso artikkelivalintaasi kohdassa **9**: **…any of ___ food I brought.**
Miten perustelet artikkelivalintasi alleviivatussa kohdassa?

---

f. Katso artikkelivalintaasi kohdassa **11**: **___ acquaintance of mine was recently…**
Miten perustelet artikkelivalintasi alleviivatussa kohdassa?

---

g. Katso artikkelivalintaasi kohdassa **15**: **…I’d like to have ___ kids.**
Miten perustelet artikkelivalintasi alleviivatussa kohdassa?

---

h. Katso artikkelivalintaasi kohdassa **15**: **…by ___ friend of mine.**
Miten perustelet artikkelivalintasi alleviivatussa kohdassa?

Kerrothan vielä **kolmen viimeisimmän englannin kurssisi arvosanasi**: ___, ___, ___

Oletko asunut yhtäjaksoisesti **yli 3 kuukautta** maassa, jossa puhutaan englantia? Kyllä___ En :