

# English Vocabulary in Early Language Learning Material

Katja Kuivalainen

Master's Thesis

Faculty of Humanities

University of Oulu

Spring 2023

## Abstract

Young learners, 1<sup>st</sup> and 2<sup>nd</sup> graders in this thesis, start learning a foreign language in the first grade in Finland. This thesis aims to determine how early language learning material promotes learning English through their vocabulary. Young learners have been required to learn a foreign language since January 2020 and teaching a foreign language to young learners is a new challenge for many primary education teachers. Thus, many teachers may rely on ready-made teaching materials such as workbooks. Young learners are a special group of foreign language learners as they are learning to read and write in one of the national languages of Finland while simultaneously learning a foreign language.

The three research questions revolve around the vocabulary content, the frequencies of the words, and how the words are recycled in the material. A corpus was formed by entering all encountered word tokens in five early language learning workbooks into a database. Each appearance of each word type was entered into the database to declare the total number of tokens in the analysed teaching material. The words of the corpus were compared to Paul Nation's word family lists that are based on the BNC and COCA word lists to define the word families and frequencies of the vocabulary.

Teaching young learners some of the vocabulary presented in the learning material was justifiable, while other words were not relevant. The frequencies of the words in the material ranged from the 1<sup>st</sup> frequency band to the 20<sup>th</sup> frequency band while some words could not be found from the BNC/COCA word lists. Some of the words presented in the material were recycled plenty of times while other words appeared only once. It can be concluded that the learning material and vocabulary both promoted and did not promote language learning as some of the word choices were less than optimal for young language learners.

Key words: EFL, frequency bands, textbook analysis, word recycling, young learners

## Tiivistelmä

Suomessa nuoret oppijat, tässä tutkielmassa 1.- ja 2.-luokkalaiset, aloittavat vieraan kielen oppimisen ensimmäisellä luokalla. Tutkielman tavoitteena on selvittää kuinka varhennetun kielen opettamisen oppimateriaalit tukevat englannin oppimista sanaston kautta. Nuoret oppijat ovat alkaneet opetella vierasta kieltä tammikuusta 2020 alkaen ja vieraan kielen opettaminen nuorille oppijoille on uusi haaste monelle luokanopettajalle. Tästä syystä monet opettajat saattavat tukeutua valmiisiin opetusmateriaaleihin, kuten tehtäväkirjoihin. Vieraan kielen oppijoina nuoret oppijat ovat erityinen joukko, koska he opettelevat samanaikaisesti lukemaan ja kirjoittamaan yhdellä Suomen virallisista kielistä ja jollain vieraalla kielellä.

Tutkielman kolme tutkimuskysymystä keskittyivät sanaston sisältöön, sanojen esiintymistiheyteen sekä sanojen uudelleen käyttöön opetusmateriaaleissa. Korpus on muodostettu syöttämällä kaikki saneet viidestä varhennetun kielenopettamisen tehtäväkirjasta tietokantaan. Jokaisen sanatyypin jokainen esiintyminen syötettiin tietokantaan, jotta saneiden kokonaismäärä analysoidussa opetusmateriaalissa selviäisi. Korpuksen sanat verrattiin Paul Nationin sanaperhelistoihin, jotka ovat muodostettu BNC- ja COCA-korpusten perusteella, jotta sanat voitiin sijoittaa sanaperheisiin ja sanaston esiintymistiheys selviäisi.

Osa oppimismateriaaleissa esiintyneestä sanastosta on perusteltua opettaa nuorille oppijoille, mutta osaa sanoista ei. Sanojen esiintymistiheys vaihteli ensimmäisestä esiintyvyysslistasta 20:nteen, kun taas osa sanoista ei löytynyt BNC- tai COCA-sanalistoista ollenkaan. Osa materiaaleissa esiintyneistä sanoista toistui useita kertoja, kun taas osa sanoista esiintyi vain kerran. Tästä voidaan päätellä, että oppimismateriaali sekä edistää että ei edistä kielen oppimista sanaston kautta, sillä osa sanavalinnoista ei ollut optimaalisia nuorille kielenoppijoille.

Avainsanat: englantia vieraana kielenä, esiintyvyysslista, nuoret oppijat, oppikirja-analyysi, sanojen uudelleen käyttö

# Table of Contents

<b>1. INTRODUCTION .....</b>	<b>1</b>
1.1 THE AIM AND RESEARCH QUESTIONS.....	5
<b>2. LITERATURE REVIEW .....</b>	<b>6</b>
2.1 YOUNG LEARNERS .....	6
2.2 THE FINNISH NATIONAL CURRICULUM OF BASIC EDUCATION 2014 .....	10
2.3 ENGLISH AS A FOREIGN LANGUAGE .....	12
2.4 ENGLISH LANGUAGE LEARNING MATERIAL .....	15
<b>3. METHODOLOGY .....</b>	<b>18</b>
3.1 MATERIAL.....	18
3.2 DATA COLLECTION TOOLS AND ANALYSES .....	21
<b>4. RESULTS .....</b>	<b>25</b>
4.1 NOUNS.....	26
4.2 PRONOUNS.....	29
4.3 VERBS.....	30
4.4 ADJECTIVES.....	32
4.5 ADVERBS .....	34
4.6 PREPOSITIONS.....	36
4.7 CONJUNCTIONS.....	37
4.8 INTERJECTIONS.....	38
4.9 DETERMINERS.....	40
<b>5. DISCUSSION .....</b>	<b>44</b>
5.1 VOCABULARY IN THE EARLY LANGUAGE LEARNING MATERIAL .....	44
5.2 FREQUENCY BANDS OF THE VOCABULARY IN THE EARLY LANGUAGE LEARNING MATERIAL.....	52
5.3 RECYCLING OF THE VOCABULARY IN THE EARLY LANGUAGE LEARNING MATERIAL .....	56
<b>6. CONCLUSION .....</b>	<b>62</b>
<b>7. REFERENCES .....</b>	<b>66</b>

## 1. Introduction

Early language learning and teaching a foreign language to young learners is a current topic in Finland as foreign language learning has been a mandatory, nationwide subject only since January 2020 (Ministry of Education and Culture, 2018). Learning a foreign language in the 1<sup>st</sup> and 2<sup>nd</sup> grade in Finland involves the challenge of students simultaneously learning the basics of reading and writing in one of the national languages – Finnish, Swedish or Sámi – of Finland. Previously, compulsory foreign language learning began in the 3<sup>rd</sup> grade (Finnish National Board of Education, 2016, p. 136) and all language learning before that is considered earlier language learning (Skinnari & Sjöberg, 2018, p. 8).

All teachers of young learners face a new challenge as they are required to teach a foreign language. The foreign language introduced for young learners has most often been English, but not always (Skinnari & Sjöberg, 2018). Even the Finnish National Core Curriculum for Basic Education 2014 [FNCC of 2014] mentions English alongside the heading “other foreign language[s]” in the foreign languages section of the curriculum (2016, p. 135). The similar trend of English being the most popular choice for an A syllabus language has continued in the 2020s when foreign language learning for young learners became mandatory.

Finnish universities prepare future primary education teachers in different ways, with some universities choosing to not offer any courses on teaching foreign languages. This means that teachers working with young learners might not have any background education on the topic of teaching foreign languages. Nikolov and Djigunović concluded their article by stating that “high-quality teacher education with appropriate focus on both language and pedagogy is an absolute must” (2019, p. 595). The lack of understanding foreign language teaching and the new initiative to teach young learners foreign languages may leave teachers to rely on ready-made teaching materials such as textbooks and workbooks. Of course, foreign language lessons do not only revolve around the use of ready-made materials, but using such materials can be justified by the fact they are produced by language and pedagogy experts. As teaching materials are produced by experts, teachers may assume that everything included in the materials is carefully considered. Nordlund mentions in her own background research that criticism has been aimed at the choices made when producing textbooks (2016, p. 48).

Therefore, this study analyses the choice of English vocabulary in all the early language learning workbooks currently available for teaching English in Finland to 1<sup>st</sup> and 2<sup>nd</sup> graders.

In this thesis *vocabulary* describes the materials' different themes and topics that words construct. *Lexicon-grammatical units* or *lexical units* combine lexicon and grammar as "vocabulary and grammar interact to the extent that there is no clear line between grammar and lexicon" (Green, 2019). Loewen and Reinders describe that *lexical units* are limited to "a specific language" and they consist of vocabulary (2011, p. 110). Brysbaert et al. (2016, p. 1) define word tokens to be "the total number of words in a corpus". "A *lemma* consists of a head word and some of its inflected and reduced forms" (Nation, 2001). In this thesis *words* are *word types* that "refer to different word forms observed in a corpus" (Brysbaert et al., 2016, p. 1). Brysbaert et al. define *word families* as "a group of lemmas that are morphologically related" (2016, p. 2). Nation broadens the term by stating that "using word families assumes that when the learner knows at least one member of the family, the other members are accessible through the application of word building rules" (2014, p. 4).

The lexicon-grammatical units gathered from the teaching material into the corpus are referred to as *words* in this study as any longer lexical units were separated into smaller sections of individual words to keep the study material similar and comparable. Single word tokens were preferred over other lexicon-grammatical units in the material. This will be the starting point for analysing the data, as comparing the corpus and Nation's word family lists will define whether the vocabulary in the young learners' workbooks is justifiably introduced in the materials.

Words tokens were recorded into the corpus in their word families under the head word. These word families are formed of head words and include all "its derived and inflected forms that can be understood by a learner without having to learn each form separately" (Bauer & Nation, 1993, p. 253). Word frequencies are discussed throughout this study and the term refers to "the number of times a linguistic item occurs" (Loewen & Reinders, 2011, p. 72). In this study, frequency describes the number of times a word token was encountered in each of the five workbooks.

Although the term young language learners can have different definitions as it can be associated for example with children aged 6 through 14 (Nikolov & Djigunović, 2019, p. 578) or children aged 4 through 12 (Rixon, 2019, p. 278), in this thesis those foreign language learning students studying in 1<sup>st</sup> and 2<sup>nd</sup> grade are referred to as *young learners*. During a government discussion on spending limits in the spring of 2018, the Finnish Government decided that foreign language learning will be added to the school curriculum with one weekly lesson per year for both 1<sup>st</sup> and 2<sup>nd</sup> grade starting in the spring term of 2020 (Ministry of Education and Culture, 2018). Even before the decision made in 2018, schools around Finland have had the option to provide foreign language classes starting from the 1<sup>st</sup> grade. Skinnari and Sjöberg's report (2018) concludes that many municipalities have offered earlier foreign language learning in the 1<sup>st</sup> grade before the nation-wide decision of 2018.

A fraction of students studying a foreign language in the first grade will have a lack of reading skills, which can be concerning as these young learners are simultaneously balancing learning a national language of Finland and a foreign language. In Finland students are not expected to know how to read or write when they begin their school journey as these skills are thoroughly practiced during the 1<sup>st</sup> grade. Starting a foreign language simultaneously means that foreign language learning cannot solely rely on reading or writing tasks, at least in the beginning of 1<sup>st</sup> grade. As reading and writing skills develop individually, teaching material must consider those young learners who will not learn to read until the end of their first school year.

Rixon has noted that while there are often numerous different languages offered for students, English is the most chosen foreign language (2019, p. 281). English is seen as a foreign language in Finland, rather than a second language, due to it not having an official status (Pietilä & Lintunen, 2014, p. 14) and thus the early language learning of English is referred to learning a foreign language in this thesis. Nikolov and Djigunović note that the learning processes of English as a foreign language or a second language have numerous similarities (2019), but the focus of this study and previous research material has been on foreign language learning settings. The FNCC of 2014 also views English as an A syllabus foreign language (2016, p. 135) which validates referring to English studies in Finland as foreign language teaching and learning.

Although early language learning is not a new topic in Finland as seen from Skinnari and Sjöberg's 2018 report, theses regarding the topic have only increased during the 2020s. There is also a lack of research in this field, especially focusing on "books intended for young learners of English as a foreign language" (Nordlund & Norberg, 2020, p. 89) and analysing the vocabulary used in textbooks (Nordlund, 2016, p. 47). This gives an incentive to study vocabulary used in workbooks that are produced to teach English as a foreign language in Finland for young learners.

Nordlund and Norberg note that many language teachers rely on the use of textbooks and workbooks in their language teaching (2020, p. 89) and with new regulations for subjects the need for teaching material to help with teaching is understandable. Weninger has also noticed how ready-made material has a central role in teaching (2018). This thesis focuses on analysing the vocabulary of five different English workbooks from two major publishers. Two of these books were published in 2018, while the other three were published in 2020 and 2021 after the decision to nationally start teaching foreign languages to young learners. This study seeks to discover how many different word tokens are encountered in the teaching material. Simultaneously, this study does not intend to find out whether the encountered words are learned by these young learners using the teaching material. The focus of this study is on vocabulary as it "is an essential part of learning a new language" (Nordlund, 2016, p. 47).

In their study of textbooks for teaching English in Sweden, Nordlund and Norberg have noted that "there is inconsistency in the choice of vocabulary included" in the teaching materials analysed in prior research (2020, p. 94). So far, no research has been published analysing similar issues in the Finnish context. Nordlund and Norberg's study is used as a starting point for one of the research questions that relies on the understanding of word frequencies. Comparing the vocabulary of five different workbooks will give a better insight to what vocabulary is taught to young learners and whether the vocabulary is similar between these five books. Working methods should take different learners and their backgrounds into consideration, and the methods used should also be appropriate for the students' age and stage of development (Finnish Board of Education, 2019, p. 28). This means that the English syllabus cannot solely revolve around only using workbooks.



For this study, a corpus was formed from the vocabulary appearing in the tasks and pictures by entering each word token onto a spreadsheet. The gathered vocabulary data was then divided accordingly to the parts of speech and further on categorised thematically. Data on the number of repetitions was also collected. As most of the vocabulary items are comprised of individual words, longer lexical units such as 'good morning' were also categorised as 'good' and 'morning' separately. After the collection of data was finished, each recorded word token was compared to the BNC/COCA word family lists to determine the frequency bands.

The numbers gathered from this study represent the number of times a student can encounter a word if the student looks at every single picture and task in the workbooks. As there is no indication of how the material is used by teachers and thus how many times a word is repeated during class or by other means of input, this study cannot take other input into consideration. Realistically teachers do not use each exercise in workbooks and therefore the reality of encountering every recorded word family as many times in a classroom setting as it was in the study is very unlikely.

### 1.1 The Aim and Research Questions

The aim of this thesis is to find out how the workbooks examined in this thesis promote English language learning through vocabulary. There are three research questions, with three sub questions for the first question, that help answer the aim of this thesis.

1. What vocabulary is introduced in the early language learning material?
  - a. How many word tokens are included?
  - b. What parts of speech are covered?
  - c. To which topics does the vocabulary belong?
2. Which word frequency bands do the words belong to?
3. How is the vocabulary recycled?

## 2. Literature review

Rixon states that changes in curricula often happen due to political or other decisions (2019, p. 279). A decree (793/2018) issued by the government amended the decree (422/2012) which enacts the Basic Education Act (628/1998) by stating that an A1 (A-syllabus) language must have at least 0,5 weekly lessons per year in the 1<sup>st</sup> grade and 1,5 weekly lessons per year in the 2<sup>nd</sup> grade, starting on the 1<sup>st</sup> of January 2020. Political decisions are not the only guiding factors of foreign language teaching as Yusupovna states that teaching goals also rely on different policies and what each society requires (2023).

The literature review section offers insight on the uniqueness of young learners, the Finnish National Core Curriculum of 2014, English as a foreign language, and English language learning material.

### 2.1 Young Learners

A common reason to begin early language learning and teaching a foreign language to young learners is the critical period hypothesis and the presumption there is an ideal age to start learning a new language (Nikolov & Djigunović, 2019). The critical period hypothesis theorises that there is a section in the developmental stage of a child when the child can acquire another language with ease, and when the period is over language acquisition becomes more difficult (Cunningham, 2011, p. 62). The sensitive period hypothesis acknowledges the possibility of learning languages after a certain age and the cut-off age is also more flexible than that of the critical period hypothesis (Loewen & Reinders, 2011, p. 155). Cunningham mentions how linguists have been baffled as children are able to learn another language when the input is very limited (2011, pp. 168–169.). This suggests that the age aspect presented in the critical and sensitive period hypotheses support the claim of effective language learning of young learners.

Significant developmental changes take effect on young learners in cognitive, social, and linguistic aspects (Butler, 2019, p. 480). When learning a foreign language, learners also go through psychological changes and one major occurring situation is making comparisons of the foreign language that is learned with the learner's mother tongue (Yusupovna, 2023, p.

522). For example, to recall vocabulary, students make connections between the language they are learning and languages they already know. Young learners cannot process information as fast as older learners, as both their memory and attention capacities are developing in a nonlinear pattern (Butler, 2019, p. 480). Thus, when working with young learners it should not be expected that when learning happens it stacks on previous knowledge. The youngest of young learners' memory capacity is often only a third of what adults have (Butler, 2019, p. 480) which supports the claim that studying should be relatively slow paced with a lot of repetition.

Butler also highlights that young learners are practicing their overall learning skills as they are learning to learn while simultaneously developing their skills in a foreign language and this learning is not always linear (2019, p. 485). As young learners' brains are developing in numerous ways, various aspects of teaching must differ from when working with older students. For instance, the older a language learner is, the more knowledge of words s/he will have acquired, and Brysbaert et al. also note that this supports the claim that older learners have not forgotten words if they have more knowledge of words (2016, p. 7). Nonetheless, recycling words in learning material and during foreign language lessons is essential.

Butler has recognised the special aspect of teaching foreign languages to young learners because their first language is evolving simultaneously (2019, p. 481). When young learners learn English, they may not understand language learning rules and thus their language learning process is like that of their first language (Nikolov & Djugunović, 2019). "Young learners are quick to learn vocabulary, slower to learn structures because words have tangible, immediate meanings whereas structures are less obviously useful" (Demircioğlu, 2010, p. 440). Teaching a foreign language should take these special aspects into consideration.

The older and more educated a language learner is, the more words they understand (Brysbaert et al., 2016) and only then should learners learn low-frequency words. Young adults know around 11,100 head words from word families (Brysbaert et al., 2016). If these words were learned linearly based on frequency bands, it would mean young adults understand words from the first 11 frequency bands, giving insight on the fact that young

learners should not be learning such words. Nikolov and Djugunović state that the frequencies of the taught vocabulary differ due to curricula (2019, p. 588). The curricula and syllabi set for young learners should take the needs and special aims of young learners into consideration. The modification for the FNCC of 2014 defines the English taught to young learners to be an A syllabus language (Finnish National Board of Education, 2019) and thus solid groundwork for future years must be done by learning appropriate vocabulary from useful frequencies.

One unclarity in the development of language is whether there is difference between monolingual and bilingual or multilingual foreign language learners (Butler, 2019). It is necessary to understand the differences between monolinguals and learners that speak more than one language while recognising the knowledge there is of young learners learning a foreign language. While undoubtable differences occur between children and their different levels of understanding, “children’s individual development in different domains may not occur in tandem” (Butler, 2019, p. 481). While studies prove that differences occur in the skill levels of monolingual students within a class, the assessment of the skills of bilingual or multilingual students is skewed because of the level set by monolingual students (Butler, 2019, p. 481). For example, the FNCC of 2014 sets objectives of instruction that should be met by each student and the objectives are aimed to be met with the amount of exposure and input offered in schools that follow the Finnish curriculum.

The basis of learning English in Finland in the 1<sup>st</sup> and 2<sup>nd</sup> grade is to help students create a positive attitude towards language learning and strengthen the students’ trust in their language skills and use of the foreign language (Finnish National Board of Education, 2019, p. 25). Helping students keep a positive attitude towards language learning creates motivation that is expected to be long-term (Nikolov & Djugunović, 2019) which not only benefits young learners struggling with their different developmental phases but also in the future as vocabulary and grammar become more difficult. Child-friendliness – implementing purposeful and engaging approaches – is a feature that should be strived for if aiming for successful language teaching (Rixon, 2019, p. 283). Yusupovna also agrees that teaching a foreign language demands understanding of its uniqueness and how to teach appropriately (2023). If a learner wants to succeed in their process of learning a foreign language it is of utmost importance to build enough knowledge of relevant vocabulary (Norberg & Nordlund,

2018). Butler describes that the English syllabi of young learners focuses on learning the ability to communicate in numerous ways and situations (2019, p. 482). Learning appropriate vocabulary in relation to young learners' skill levels and their age is essential when supporting their path to becoming better communicators in the targeted foreign language.

Young learners are capable of doing tasks together with others and interested in stories (Butler, 2019, p. 480). As stories are a general interest of young learners, the use of them in teaching vocabulary can certainly be justified and simultaneously it supports the objectives of the FNCC of 2014 introduced in section 2.2. Incidental learning is considered widely to be the best way to acquire and learn vocabulary, as large numbers of vocabulary are not learned only by intentional learning (Hulstijn, 2012, p. 362). Yusupovna describes foreign language lessons as settings for intercultural communication, and those situations require a lot of resilience and perseverance from the participants (2023, p. 521). The process of learning vocabulary should not be hurried, especially with young learners, and practicing communication with known words is a key to mastering the target language.

If assessment of young learners must take the learners' "unique age-related and environmental characteristics" into consideration (Butler, 2019, p. 478), teaching young learners and choosing appropriate content for them must also be an essential part of curricula and syllabi creation. Using word family lists to define the knowledge level of a learner is problematic as Nation describes the key element being learner knowledge, and the understanding of word family lists may be incomprehensible to learners (2021, p. 696). A prime example is the verb "be" and the different declinations that do not look alike. Language learners may not understand different word forms belong to the same head word, thus using word families to gain understanding of young learners may not always be beneficial. Bauer and Nation have realised that a learner's understanding of word families grows when their "knowledge of affixation develops" and the learner understands what head word ties different word types to that word family (1993, p. 253).

Young learners are a special group to teach a foreign language as earlier has been concluded. Bland describes young learners aim to learn through holistic means and that these learners are very keen on understanding the world around them, and this provides a great opportunity

to expand young learners' knowledge of their surroundings in multiple ways (2015). The next section will discuss the Finnish National Core Curriculum of Basic Education 2014 and how these exceptional learners and conditions are considered in it.

## 2.2 The Finnish National Curriculum of Basic Education 2014

The “whole educational program, comprising all subjects taught” is referred to as a curriculum, while syllabus defines “the documented structure and learning components for a course of study in a particular subject” (Rixon, 2019, p. 278). As the subject surrounding this study is English, the guidelines the FNCC of 2014 has defined for the subject is referred to as the syllabus of English in this thesis. The design of the language curriculum and syllabus for young learners must be different than that of older language learners, as young learners need various child-friendly approaches embedded in their learning (Rixon, 2019, p. 278). The FNCC of 2014 resolves this by implementing a separate curriculum that only applies to young learners in 1<sup>st</sup> and 2<sup>nd</sup> grade. Curricula for young learners relies on information set by ministries gathered from both international and national decisions (Nikolov & Djugunović, 2019).

The additional regulation concerning the curriculum of 1<sup>st</sup> and 2<sup>nd</sup> grade of the FNCC of 2014 has set objectives of instruction for each subject and the foreign languages syllabus for young learners focuses on “growing into cultural diversity and language awareness”, “language-learning skills”, “evolving language proficiency and interaction, text interpretation, text production skills” (Finnish National Board of Education, 2019, pp. 26–28). The FNCC sets the objectives of instruction for all subjects including foreign languages, after which each teacher is free to plan their syllabus in a way that the objectives are met.

Numerous countries have decided “to include objectives that go beyond linguistic categories” (Rixon, 2019, s. 285) and Finland is no exception to this trend. The FNCC of 2014 includes seven different transversal competences that are to be included in all teaching throughout the school year and in each subject. The aim is to develop learners' knowledge on a wider scale (Rixon, 2019) rather than setting learning objectives only on the curricula or syllabi of subjects taught at school. Learning a language can help students develop their transversal competences for example in *Cultural competence, interaction and self-expression* (T2), and

*Multiliteracy* (T4) (Finnish National Board of Education, 2016, pp. 22–23), which supports Rixon’s claim of aiming to reach a wider knowledge level.

Because the FNCC of 2014 encourages transversal competence, supporting young learners as they are learning to learn is fundamental on their path to becoming better foreign language learners and language users. Rixon mentions the possible forming of motivation as one of the positive reasons to include meaningful ways of teaching when working with young learners, but she also describes the difficulty of designing curricula based on meaningful working methods (2019, p. 284). This does not concern Finland in a similar way, as the FNCC in place only guides and obligates teachers to some aspect, but the way they choose to approach objectives of instruction and teach can be done in any appropriate way imaginable.

The new addition to the FNCC of 2014 concerning foreign language teaching in 1<sup>st</sup> and 2<sup>nd</sup> grade states that topics covered in class should be chosen together with the class to ensure that other guidelines such as the joy of learning is guaranteed while learning a foreign language (Finnish National Board of Education, 2019, p. 25). Young learners’ foreign language lessons focus on other things rather than learning the grammatical aspects of a language or its lexis (Rixon, 2019, p. 284) and thus taking the students’ interests into consideration makes learning a foreign language and its vocabulary more enjoyable. Demircioğlu highlights the importance of providing enthusiastic ways of learning vocabulary to engage and motivate students to learn a foreign language (2010).

Generally, when teaching a language, the syllabus should regard its contents from the viewpoint of the whole curriculum (Rixon, 2019, p. 278). Understanding lexical threshold provides information which should be used on deciding appropriate vocabulary for learners to aim for (Laufer & Ravenhorst-Kalovski, 2010). Rixon mentions that learning and curricula have shifted to a more learner-centred way of working as learners practice working more autonomously and focus on developing their learning skills (2019). The FNCC of 2014 mentions similar skills as transversal competences of which *Thinking and learning to learn* (T1) supports life-long learning and students viewing themselves as learners who can observe the way they learn (Finnish National Board of Education, 2016, p. 21).

Rixon mentions the idea of an organising principle “that determines the selection, grouping together, and sequencing of what is taught” (2019, p. 283). She continues that the curricula designed for young learners does not generally pinpoint “development of language skills and knowledge of language items” as an organising principle (2019, p. 283) and the trend can be seen in teaching material. Rixon has concluded that more child-friendly activities such as stories and games are more likely to be the organising principles as they are viewed more meaningful because of their close relatedness to the lives of children (2019). If the objectives of instruction for foreign languages are met, teaching through topics or themes, or by any other means, can be successful especially if the students’ interests are taken into consideration.

The FNCC of 2014 offers the guidelines and support for teachers of young learners, but its flexibility and the trust that teachers know their students is what provides the best opportunities for young learners to learn a foreign language. Curricula and syllabi must not only consider the speciality of young learners, but also the fact English is the first foreign language for many students. The aspect of English being a foreign language is reviewed in the next section.

### 2.3 English as a Foreign Language

Globally, English is primarily chosen as a foreign language subject in primary schools (Rixon, 2019) which is the case in Finland too (Skinnari & Sjöberg, 2018). English is most often chosen in the place of other languages, including local languages (Rixon, 2019, p. 281). As Swedish, one of the national languages of Finland, is already a compulsory school subject starting latest in grade 6, choosing Swedish may be overlooked as the first foreign language as it will be learned eventually. Rixon reminds that English is rarely the only foreign language choice given to students (2019, p. 280). From Skinnari and Sjöberg’s report it is clear that municipalities offer up to six or even seven different choices for the first A syllabus language such as Chinese, English, Estonian, French, German, Russian, Spanish, and Swedish (2018). Many EU countries have various national and minority languages, and learning more than one foreign language at a primary school level is not an unfamiliar setting (Rixon, 2019, p. 280).



Rixon refers to time that has been dedicated to English as a subject in different curricula as “relatively modest” (2019, p. 282). While that is also true in Finland, the weekly hours for each subject are the bare minimum, not hours that cannot be exceeded. Learning settings should also be diverse to support learning with minimal teaching hours (Butler, 2019, p. 479). The FNCC of 2014 describes the conception of learning to consist of plenty of interactions with others and multiple environments where students can learn (Finnish National Board of Education, 2016, p. 17).

An overall understanding of vocabulary is a major factor in further language skills of literacy and reading comprehension (Nordlund & Norberg, 2020). In their study, Nordlund and Norberg continue to state that “a larger vocabulary will, thus, lead to better lexical coverage of both spoken and written texts in general”, but the amount of necessary lexical coverage is dependent on the text (2020, p. 91). Learners of English as a foreign language are likely to have various end goals they are working towards and thus their learning needs to differ from students learning English as a second language (Butler, 2019, p. 479). English is a foreign language in Finland as it has no official status in the country. Butler continues to critique the assumption of different end goals as it may be too generalised (2019, p. 479), but the first aim of young learners is to learn vocabulary, which is a key element of language learning in a foreign language setting (Azim et al., 2020). Barclay and Schmitt advise turning to word lists when unsure of what vocabulary should be taught (2019, p. 803).

The goal of learning every single word in a target language is not a practical goal for a second language learner (Nation, 2006) thus starting the language learning process from the most frequently occurring words is reasonable. Learning frequently appearing words is important for all learners (Barclay & Schmitt, 2019) as words that appear less in text provide smaller coverage (Laufer & Ravenhorst-Kolovski, 2010, p. 24). While Nation refers to second language learners, obviously the situation is the same with foreign language learners.

Yusupovna has realised that foreign language teaching differs from the learning of native and second languages as the setting when learning a foreign language is often an artificial one rather than an authentic one (2023, p. 519). While English can be heard and seen around many parts of Finland and on the internet, Yusupovna believes most of the foreign language

communication between students happens in a class setting (2023, p. 519). The classroom is a safe place to learn vocabulary and practice it, even though the setting may be artificially provided. The artificial setting is one of the differences that separates English as a foreign language from a second language and thus giving the assumption that students are exposed to the language differently (Butler, 2019, p. 479).

Paul Nation states in his study (2014) that repetition is a necessary part of learning words, and encountering words regularly enhances the possibilities of learning them. Encountering new words and thus broadening known vocabulary is not enough as words also need to be encountered frequently enough for notable acquisition (Nordlund & Norberg, 2020). In their study, Nordlund and Norberg emphasise on the importance of repeating words to learn them and develop vocabulary (2020). Nation has defined a minimum “safe goal of 12 repetitions” in his 2014 study as it offers the language learner to encounter words in different contexts and the possibility for “several unassisted retrievals” (p. 3). Norberg and Nordlund have not concluded on the amount of repetition and how many times a word should be encountered in textbooks so it is guaranteed that a word is learned (2018, p. 464). While learning a word is not guaranteed, young learners need the opportunity to encounter the word enough times that learning could be possible through repetition and recycling.

Native speakers acquire approximately 1,000 word families annually until around the age of 20, while unpublished research at the time estimates that based on a vocabulary size test on English as a foreign language learners, who are completing advanced studies, have a receptive vocabulary of approximately “8,000 to 9,000 word families” (Nation, 2006, p. 60). Therefore, young learners are not required to know vocabulary from such frequency bands. Nation and Waring present that a vocabulary size of 2,000 known word families amounts to a text coverage of 79,7% which is a coverage of more than 3/4 (1997). Laufer & Ravenhorst-Kalovski concluded in their study that optimal knowledge of word families is met when a learner knows 8,000 word families as this provides a 98% covered, but also note that a 95% coverage is met with a threshold of 4,000 to 5,000 word families (2010). This is something foreign language learners can achieve, but not something that should be aimed for in the first years of learning a foreign language.

While some of the recommendations for learning English as a foreign language are undecided, learning happens in artificial settings and with the help of vocabulary. Whether English language learning material considers the young learners' uniqueness in learning a foreign language, the guidelines set by curricula and syllabi, and the speciality of learning English as a foreign language is reviewed in the following section.

#### 2.4 English Language Learning Material

Rixon concluded in her article that forming curricula and syllabi is a team effort where those who are working on designing said material should work with those who use the material and vice versa (2019, p. 291). Just like any other subject in curricula the capabilities and needs of each student differ more strongly from one another than before (Butler, 2019, p. 284). Luckily teaching material often include tasks of various difficulty helping teachers to diversify tasks offered to students. The need to differentiate tasks may be a result of young learners encountering English earlier than before while others start with the basics at school. Media offers language immersion outside the formal school setting (Butler, 2019, p. 479), which may affect the learning materials used in school if publishers expect prior knowledge of the target language.

The topics discussed and dealt with in class are to be chosen by the students themselves of topics that interest them, but at the same time the topics taught should vary (Finnish National Board of Education, 2016). Norberg and Nordlund discovered in their study of English learning material in Sweden that the topics of colours, "words for pets and family members, and ways of expressing likes" are viewed as important themes by publishers (2018, p. 469). The results and discussion of this thesis will comment whether similar themes are seen in the teaching material in Finland and what frequency bands the vocabulary is found in.

Norberg and Nordlund (2018) share that the most common words in discourse are the first 1,000 to 2,000 words meaning the words are from the 1<sup>st</sup> and 2<sup>nd</sup> frequency bands. Young learners are presumably interested in vocabulary out of the scope of those frequency bands, especially when taking a cultural aspect into consideration. Brysbaert et al. recall that most previous studies on understanding what language learners know have focused on the number

of lemmas, that is a categorisation type that does not count proper nouns into data and fixes errors in spelling (2016).

In her article, Rixon states that when changes in curricula happen analysing whether teaching material should be changed or reimagined is necessary (2019, p 268). As some of the young learners are encountering English for the very first time when they start their A syllabus language at school it is essential to start with the basics. The basic teaching material, that studying a foreign language should start with, should be filled with the first 1,000 words as they are the most commonly occurring words in English (Nation, 1993). Thus, the teaching of young learners should focus on the most commonly occurring words so that their vocabulary knowledge builds in a necessary way and a “solid basis for further learning” is guaranteed (Norberg & Nordlund, 2018, p. 464).

However, Norberg and Nordlund understand that gaining knowledge of vocabulary by studying words does not guarantee fast access to a language (2018). Azim et al. have understood that “recurrence, repetition and recycling” are all important when listing ways that provide positive outcomes of vocabulary learning (2020, p. 93). Thus, considering what vocabulary is taught is not enough but considering the number of times words appears in teaching material is also essential. Previous research has analysed recycling of words, but not thoroughly enough in different authentic materials for children (Gardner, 2008). Nordlund and Norberg have studied the recycling of vocabulary in teaching material in English books published in Sweden (2018; 2020). Gardner states that research has been especially interested in the recycling of words that are not the most common, highest frequency words such as articles and most commonly needed verbs (2008, p. 93). Nation has gathered information from studies that support the claim that the most common words in English are learned very well (1993, p. 193). Nonetheless this does not happen without enough repetition and recycling of words.

Nordlund & Norberg’s 2020 study concludes that new words are introduced and there is a lack of word recycling, especially of words that are categorised as low frequency. Their study of 2018 resulted in understanding that young learners’ English teaching material in Sweden offer different amounts of exposure to vocabulary, and that vocabulary in general is not

chosen accordingly (Norberg & Nordlund). Norberg and Nordlund also state that not enough research has been focusing on the vocabulary of textbooks (2018, p. 463) which suggests that research and studies on teaching material is highly needed.

There clearly is need for studies of vocabulary in language learning material if the outcomes in Sweden have been of this calibre. Barclay and Schmitt state that defining the vocabulary to teach to young learners is simple as they need to learn words that help them navigate through their day such as classroom related vocabulary (2019). After the methodology aspects of this thesis are discussed in the next section, the results and discussion will clarify whether the Finnish material used to teach English for young learners has realised the aspect of vocabulary teaching.

### 3. Methodology

In this study a corpus was formed by gathering words from five early English learning workbooks used in the 1<sup>st</sup> and 2<sup>nd</sup> grade in Finland with young learners. Analysing the material is achieved with textbook analysis, where linguistic focus is on a textbook and its contents (Weninger, 2018). First the selection criteria are outlined, and a general overview of the material is provided. After this, the formation of the corpus and the categorisation systems are introduced along other data relating to data collection. Lastly, the data analyses process is discussed.

#### 3.1 Material

Choosing the material for this thesis was easy as the five books included in the study are the only early language learning workbooks for English that are currently available as printed books, and because of that there is little variation in the publishers of the workbooks. The material of this study is limited to physical books which comprise the material young learners will work with both in class and at home. Thus, every single commercially available workbook for Finnish-speaking teachers of young learners has been included and examined in this study.

SanomaPro and Otava are two major publishers for teaching materials and are the ones that have provided the earliest material for early language learning, as expected. Additionally, Otava is also working on a new series for grades 1 to 6, *Skylight*, and the first books will be published in 2024 (Otava, 2023). In addition to these materials, another publisher, Edukustannus, offers an online material (*Get Ready! 1–2*) for teaching early English for young learners, but currently the publisher does not have a physical workbook to go along with their online material (Edukustannus, 2023). While the workbooks were published by two publishers, the material is produced by five pairs or groups of different people.

The five different early language learning books that are analysed in this study are introduced in Table 1. Two of the workbooks, *Jump in!* (Haukka & Rantanen) and *Go!* (Kanervo & Laukkarinen), were published in 2016, before the decision to include early language learning in the national curriculum was in place. The other three books, *Come with me! 1* (Harjula, Heikkinen, Lyykkö, Pere, Ryösä & Turpeinen, 2020), *Come with me! 2* (2021) and *High five! 1–*

2 (Kalaja, Korpela, Kuja-Kyyny-Pajula, Mäkinen & Pelli-Kouvo, 2020), were published in the 2020s after the decision was made in 2018. Four of the workbook titles include a verb that is presented in the imperative form followed by an exclamation mark, and *High five! 1–2* also uses an exclamation mark as “high five” can be considered as an interjection.

Book	Publisher	Year of Publication
Come with me! 1	SanomaPro	2021 (1 <sup>st</sup> – 3 <sup>rd</sup> edition)
Come with me! 2	SanomaPro	2020 (1 <sup>st</sup> edition)
Go!	SanomaPro	2016 (1 <sup>st</sup> edition)
High five! 1–2	Otava	2020 (1 <sup>st</sup> edition)
Jump in!	Otava	2016 (1 <sup>st</sup> edition)

**Table 1.** The teaching material.

*Come with me! 1* is the first book of the *Come with me!* series and has 88 pages. There are eight chapters and an introductory chapter at the beginning of the workbook. Each of the eight main chapters follow the same structure. First, new vocabulary is introduced, after which there are tasks that sometimes include text. Each chapter includes a dialogue, that is in text form, which relates to a listening task, and each chapter includes an interactive game that focuses on speech. At the end of each chapter there are tasks that bind together what is learned in the earlier exercises of that chapter and a self-evaluation. The second book in the series, *Come with me! 2*, follows the same pattern. The series’ second workbook has four introductory chapters and eight main chapters in its 96 pages. Both workbooks have extra tasks at the end of the book for each main chapter. The *Come with me!* series has a page at the end of its books where the young learner can write words, but the workbook does not disclose what types of words the page is aimed at.

*Go!* has 96 pages and 15 chapters, the last one partly having exercises that summarise the vocabulary from the preceding 14 chapters. Like the *Come with me!* series, *Go!* proceeds in a continuous format in each chapter and similarly starts off by introducing the new vocabulary. The core vocabulary is followed by a song with its lyrics visible in the workbook and exercises that include listening, conversation, writing, and a game. *Go!* also ends each chapter with a self-evaluation task.

There are 12 main chapters and one introductory chapter in the 96 pages of *High five! 1–2*. Each main chapter has extra tasks at the end of the book and a section that introduces the four seasons. The chapters of *High five! 1–2* also follow a common structure, starting with pictures that lead into the topic of the chapter, followed by the new core vocabulary. There are several listening tasks and an exercise that requires drawing. There are no exercises in the workbook that directly require writing. Pictures and games are included in each chapter of the workbook.

*Jump in!* is the fourth workbook with 96 pages. Once again, there are three introductory chapters before then ten main chapters. There are extra tasks at the end of the book for each chapter and lists where the young learners can list words with certain phonemes (e.g., [i] versus [i:]) from each chapter. Just like the other materials, *Jump in!* also follows a certain order in each of the chapters. Ten new words are presented at the bottom of the first two pages of each chapter. Each chapter also introduces functional tasks and games, and writing, listening, and speaking exercises. Some of the chapters also include song lyrics and drawing exercises.

The *Come with me!* series offers a separate book for both 1<sup>st</sup> and 2<sup>nd</sup> grade, while the rest of the workbooks are designed to be used in both or either of the grades. As there are weekly English lessons both in 1<sup>st</sup> and 2<sup>nd</sup> grade, the other series' books cannot be the only material used to teach English as there is not enough material to cover both school years. Each workbook is set up in a similar way – new core vocabulary is introduced in the beginning of the chapter and numerous tasks that require discussion and speaking English are included. The workbooks include listening tasks with the dialogue, or other form of text for the students to see, and games and other functional activities. The *Come with me!* series includes an exercise in each main chapter that introduces other foreign languages by making connections to the English language in various ways. *High five! 1–2* and *Jump in!* do not include a self-evaluation task at the end of any of the chapters for the student to fill out unlike the other three workbooks.



### 3.2 Data Collection Tools and Analyses

In this thesis, word forms are referred to as *word types*, which are either explicitly different words, different forms of the same word (e.g., singular and plural), different words that belong to the same word family, or word forms that are the same while spelled differently in regards of capitalisation (Brysbaert et al., 2016). Each word type is recorded as they offer insight on the total number of word tokens and different words types young learners encounter when working with the early language learning material of this study. The definition of *word* in this study considers individual words of the material, but not longer lexical units. Most of the word types were recorded into the database as they were found in the material, but “let’s” was recorded as “let” and “us” and “don’t” as “do” and “not”. The workbooks included tasks that require writing and if there was only one answer to the exercise (e.g., the answer “book” in a task with a picture of a book) it was recorded into the database assuming that the tasks are done accordingly.

Studies that conclude that learners lack morphological-awareness and thus are not able to create connections between words from the same word family have focused on native speakers of English (Gardner, 2008, p. 99). Young learners who are learning a foreign language should not be expected to understand the morphological system better than native learners. While the words that were entered into the database were categorised by the word family under the head word, each separate word type was also recorded. The word types were entered into the database along with information relating to which part of speech each word belongs to, in which chapters the word is encountered in and how many times, and which frequency band it belongs to when comparing the head word to the BNC/COCA word family lists. The categorisation by word families enabled the comparison of the number of word families in the different workbooks, which provides different information than the number of different tokens encountered. Nouns are also categorised by themes and verbs by their base form.

Each word type was entered into the database manually and some word types required modification. Each word type is recorded under the head word of the word family, which is gathered from the BNC/COCA word family lists. Different word types, such as singulars and plurals or different verb conjugations, were recorded under the head word, but counted

separately towards the total number of word tokens. There are words that were categorised into two sections as sometimes a word (family) was encountered in different contexts, such as “orange” as an adjective (the colour) and a noun (the fruit). Different word types also included different spelling of the same words (e.g., “grey” and “gray”) and synonyms (e.g., “grandpa” and “granddad”), which can for one be seen as a result of choosing different varieties of English in the material.

After the corpus was formed, the items were compared to Nation’s word family lists. The lists were made with the aim to provide a list of useful words to learn (Nation, 2021) and entering the word types from the material into the database with the word family of each token helps answer the research questions. The family lists used in this thesis are formed from the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA). These family lists are formed by the frequencies and range of words, where the 1<sup>st</sup> 1000 family words include the most common 1,000 words (Nation, 2014). Nation describes the BNC to include “British, adult, formal, informative language” (2004, p. 3). Using the word family lists will provide the information of which words belong to the same root family and thus are from the same frequency band. Moreover, categorisation within each part of speech is easier and the analysis of the results is clearer as the database is structured logically. Including both a British and an American corpus provides a range of words that could have been missed if the word family lists only used one of the two.

The five workbooks analysed in this thesis present the different exercises either with reoccurring pictures that relate to the type of exercise (e.g., headphones for a listening task in the *Come with me!* series), or the task is explained and instructed in Finnish. While the workbook used for this study of *High five! 1–2* had instructions in English, the series also offers another version of the workbook with Finnish instructions, which is most likely the version used in Finnish-speaking classes. Therefore, the English instructions in *High five! 1–2* are not analysed when gathering the vocabulary from the material. Doing this provides the opportunity to put each workbook in more of an equal comparison, as the other four workbooks have their instructions in Finnish, and it would skew the results. Thus, the material analysed in this thesis consists of the vocabulary introduced in the tasks themselves, names of tasks (e.g., “Slap the card!” in *Jump in!* where the Finnish translation was also present), and

in the pictures seen in the books. All the workbooks included listening tasks, but these were left outside the scope of this thesis unless the listening exercise was also presented in text form in the book. In addition, the vocabulary introduced on the front and back covers of the books was also recorded into the database as books included vocabulary such as colours and days of the week on them.

First, each book was skimmed to provide an overview of the different types of workbooks, after which all the word types were entered into a database where the words from the material were distributed into different sections based on the parts of speech they belong to. Noun forms were further divided into themes that the words are associated with in the material and verb types were specified by adding the detail of their base form. Each workbook introduced new vocabulary under a different theme that the words related to, which helped dividing the nouns into different categories. Dividing the material into different topics helped analysing the vocabulary, as the vocabulary taught should be relevant (Nordlund & Norberg, 2020, p. 90) for the age group and their skills, and the division will also help answer the research questions.

Longer lexical units, such as “good morning”, “thank you”, “pencil cases”, and “corn flakes” were entered into the database as individual words. Longer lexicon-grammatical units appeared mainly in speaking exercises, listening tasks, and songs. Some of the lexical units with two words were a part of the new vocabulary taught in a chapter (e.g., “I’m fine, thanks”), but mainly they were single words, which made the entering of word types into the database easier. Independent words made up most of the lexical units in the material, therefore individual words were entered into the database and the corpus was formed based on them. Some of the lexical units in the material could not be recorded into the database as they were not proper words. The third main chapter in *High five! 1–2* includes a picture of a scene from a classroom with a world map on the wall that has the text “MERICA” which is missing an ‘A’ as it is under a USA flag. These types of words were not recorded into the database as it should not be assumed that young learners know the word that is missing letters. While the BNC/COCA word family lists do not include proper nouns, they were entered into the database nonetheless because they were included in the total number of word tokens as they are actual words.

Nation notes that the formation of word lists often happens for the design of syllabi and to discover what is necessary to learn (2004, p. 3). Each of the authors of these workbooks have chosen the vocabulary in the material by their own reasoning and it will remain unknown whether data relating to word frequencies has been considered in the process. Nonetheless, this thesis will compare the formed corpus to the BNC/COCA word family lists to analyse the results and find out what type of vocabulary is introduced in the material and whether it supports language learning. Comparing the data to Nation's word family lists is not unproblematic as they are often made based on a corpus that is designed from adult language (Nation, 2004, p. 3). While this study focuses on the teaching material of young learners and not on material designed for adults, the words in the material should occur in the first few frequency bands and such common words are available in the BNC/COCA word family lists. The BNC/COCA lists offer enough information about the frequencies and word families to justify using them in the scope of this study.

## 4. Results

The total number of word tokens that were relevant to this study was 9,377. This is the total number of different words appearing in different word types and their total number of repetitions in the material. The word tokens of the material were entered into the database under the head word, thus giving the number of word families in the material. There were 1,188 different word families and the few overlapping words in different parts of speech are excluded from the total number. Consequently, the results of this thesis are presented in nine sections that each represent a part of speech. To analyse these results each encountered word token was linked to a word family so as to answer the research questions related to the frequency bands. Table 2 presents the material with information of word families and the total number of word tokens for each workbook.

<b>Book</b>	<b>Word families</b>	<b>Total number of word tokens</b>
Come with me! 1	180	1,242
Come with me! 2	227	1,561
Go!	263	2,654
High five! 1–2	234	1,427
Jump in!	286	2,493

**Table 2.** Number of word families and tokens.

The vocabulary gathered from the teaching material appeared either in only one of the books, in a two to four of the books, or in each of the five books. Each workbook also introduced words that were either repeated regularly throughout the whole book or appeared only once or twice. For example, colours and numbers were repeated regularly throughout the material while other words, such as the verb “walk” in both *Jump in!* and *High five! 1–2*, appeared only once or twice in the whole book respectively.

The recorded vocabulary was tagged into the corpus based on the part of speech each word belongs to, which helped both categorise words in a logical way and give a better understanding of what type of vocabulary is introduced in the young learners’ learning material. As there were more nouns than any other word types in the corpus, the greater variety of the nouns in all five workbooks was not surprising. There were lexical units in the

material that can be viewed as synonyms of one another, which mostly led to those words being from different word families and thus there being the chance that they belonged to different frequency bands. An example of synonyms that were from different frequency bands was *Jump in!* and *High five! 1–2* using the word “rubber” from the 4<sup>th</sup> frequency band while *Go!* and *Come with me! 2* used the word “eraser” from the 5<sup>th</sup> frequency band.

The upcoming subsections will present the results in nine different parts of speech that the corpus’ words were categorised as: nouns, pronouns, verbs, adjectives, adverbs, prepositions, conjunctions, interjections, and determiners. The latter contains three main sections, including articles, numbers, and possessive determiners. Sometimes determiners are categorised belonging into the other parts of speech, but because of the diversity of the category determiner became its own section.

#### 4.1 Nouns

The part of speech that appeared undoubtedly the most in each of the five chosen workbooks was nouns. Most of the nouns that were gathered from the teaching material were easily grouped into different themes that they were presented alongside with in the five workbooks. Some words, such as “noise” and “safety”, were not easily grouped under a clear theme leaving such words to be tagged under “miscellaneous”. Some of the nouns gathered from the material could not be sectioned into a frequency band because “in general, lemmas exclude proper nouns” (Brysbaert et al., 2016, p. 2) and the material presented proper nouns such as “English”, “MacDonald”, and “North America” that were found in pictures and songs for example. There were 352 different noun word families introduced in the teaching material. Nouns presented both in the singular form and plural form.

<b>Book</b>	<b>Word families</b>	<b>Total number of word tokens</b>
Come with me! 1	90	413
Come with me! 2	102	623
Go!	127	851
High five! 1–2	124	460
Jump in!	167	883

**Table 3.** Nouns.

Table 3 presents the number of word families and total number of nouns in each of the five workbooks. The themes that the nouns were tagged as are presented in Table 4 below. To help understand the vocabulary of each topic a few examples are provided from each theme.

GROUP NAME	EXAMPLES	GROUP NAME	EXAMPLES	GROUP NAME	EXAMPLES
Animals and pets	<i>cow, dog</i>	House	<i>bed, table</i>	Proper nouns	<i>Asia, Euro</i>
Body parts	<i>head, shoulder</i>	Items	<i>camera, suitcase</i>	School	<i>classroom, rubber</i>
Clothes	<i>jumper, shoes</i>	Jobs	<i>detective, spy</i>	Seasons	<i>spring, winter</i>
Family and friends	<i>grandad, mum</i>	Miscellaneous	<i>alien, corner</i>	Time	<i>afternoon, morning</i>
Food	<i>carrots, orange</i>	Nature and outside	<i>grass, sun</i>	Transportation	<i>bus, wipers</i>
Head names	<i>colours, time</i>	Onomatopoeia	<i>oink, swish</i>	Toys	<i>games, toy</i>
Hobbies	<i>football, piano</i>	Places	<i>bank, farm</i>	Weekday	<i>Monday, Tue</i>

**Table 4.** Tagged noun groups.

Only five nouns appeared more than 50 times in the whole material. The most common noun was “name” with its 71 appearances, followed by “corner” (56), “play” (56), “school” (55), and “football” (53). 98 different word families were presented only once in the whole material meaning that 27,76% of the 353 different word families appeared just once. The number of appearances of individual words in the whole material ranged from one to 71 and the average number of appearances per noun was 9,52. These aspects along with the details of frequency bands are further processed in the discussion section of this thesis.

*Jump in!* contained the most distinct word families (167) and the total number of nouns in the two different forms was 203. The total number of appearances was also the highest of nouns in the material being 883. The noun “colour” appeared 22 times in the workbook making it the most common noun. “Number” appeared 16 times, but all the other 165 word families appeared 15 times or less. The nouns belonged mainly to the 1<sup>st</sup> and 2<sup>nd</sup> frequency band, but the highest frequency band word in the whole material is “oink” from *Jump in!* and it belongs to the 20<sup>th</sup> frequency band.

*Go!* introduced 127 word families in 138 different forms, meaning that under 10% of the word families included both the singular and plural forms of the nouns. “School” appeared 34 times in the material, which was the only noun to reach more than 30 repetitions in *Go!*. The

frequency bands of the nouns range from the 1<sup>st</sup> (e.g., “queen”) to the 15<sup>th</sup> (e.g., “hopscotch”). The total number of appearances of nouns was 851.

There were 124 word families in *High five! 1–2* and altogether 145 different word types. *High five! 1–2* included the most nouns that could not be categorised into a frequency band as they were proper nouns such as the names of continents. The most common noun was “school” with its 13 appearances, which amounts to only 2,83% of all the appearances (460). *High five! 1–2* chose to introduce the word “hoodie” which belongs to the 18<sup>th</sup> frequency band, but most of the words were from the 1<sup>st</sup> and 2<sup>nd</sup> frequency bands.

*Come with me! 1* introduced the least (90) word families and also the least number of appearances (413) throughout the whole workbook. The word families presented in 103 different forms with some appearing only in either the singular or plural form and very few in both forms. “Play” and “corner” were the most common words in the workbook because the nouns appeared in each chapter as a repeating section. While most words were from the 1<sup>st</sup> and 2<sup>nd</sup> frequency band, there was a great deal of variety as nouns varied from the 3<sup>rd</sup> through to the 10<sup>th</sup>, 14<sup>th</sup>, and 18<sup>th</sup> frequency bands. 44,44% of the nouns in *Come with me! 1* appear only once in the whole workbook.

A total of 102 word families were introduced in *Come with me! 2* and they were very different from the nouns in the series’ first book, and thus young learners are introduced to new nouns in the 2<sup>nd</sup> grade. There were 130 different word types in total as many of the main nouns taught in each chapter appeared both in the singular and plural form at some point in the workbook. Like the nouns in *Come with me! 1* the series’ second book included nouns from a great variety of frequency bands ranging from the 1<sup>st</sup> to the 13<sup>th</sup>. Because of the way the series is built, “play” and “corner” were also the most common nouns in this workbook as they each appeared 30 times. Other than those two nouns, the word family “Monday” was introduced the most with 18 encounters of its two nouns “Monday” and “Mon”.



## 4.2 Pronouns

There were 15 different words in the study material that were categorised as pronouns. These 15 different words consisted of 13 different word families. The pronouns were tagged as personal pronouns (“I” and “me”, “you”, “we” and “us”, “he”, “she”, “it”, “them”), relative pronouns (“that”, “who”), demonstrative pronouns (“this”, “that”), interrogative pronouns (“what”, “where”) and indefinite pronouns (“something”). Each pronoun belongs to the 1<sup>st</sup> frequency band and thus all pronouns in this study are in the top 1,000 most commonly occurring English words. The total number of word tokens was 772.

Out of the five workbooks *Come with me! 2* had the most different pronouns (10) and *Jump in!* had the least (6). *Go!* had eight pronouns, which was the median of the five workbooks, but the eight pronouns appeared a total of 228 times throughout the workbook, which was noticeably the most of any of the five workbooks. On the other hand, *High five! 1–2* had the second most pronouns (9) of the material, but it was the only workbook that did not reach triple digits in the number of repetitions as the nine pronouns appeared only 77 times. Table 5 presents the number of word families and total number of pronouns in each of the five workbooks.

Book	Word families	Total number of word tokens
Come with me! 1	7	131
Come with me! 2	10	172
Go!	8	228
High five! 1–2	9	77
Jump in!	6	164

**Table 5.** Pronouns.

The most reoccurring pronoun word family was “I” which includes the personal pronouns “I” and “me”. Even though the word family “I” also includes “my”, it is categorised as a determiner in this study. While “I” was the pronoun that appeared the most there was great variety in its appearance as *Come with me! 2* included 99 “I” or “me” while those pronouns were only found 25 times in *High five! 1–2*. “I” evidently belongs to the 1<sup>st</sup> frequency band. The second most reoccurring word family was “you”. Similarly, to the word family “I”, the

word family “you” includes the lexical unit “your”, but in this study “your” is dealt with as a determiner. “You” belongs to the 1<sup>st</sup> frequency band and the most it appeared in any of the workbooks was 83 times in *Go!* and the least in *Come with me! 2* (26).

There was a significant drop in the number of occurrences of other pronouns in all of the material as “I” and “you” word families consisted of 69,52% of the lexical units categorised as pronouns. There is variety within the material concerning the least common pronoun. In *Jump in!* the least common pronoun was “he” (2 appearances), in *High five! 1–2* it was “where” (1 appearance), in *Go!* it was “that” (1 appearance), in *Come with me! 1* it was “this” (1 appearance), and in *Come with me! 2* it was “something” (1 appearance).

#### 4.3 Verbs

There were 76 different word families of verbs in the material, but 93 different verb forms. The total number of word tokens that were verbs was 1,659. Most of the verbs appeared in the base form, while others appeared either in the third person singular, past participle, continuous present, and past simple. The verb “let” never appeared as such in the material, but rather as “let’s”, but to keep the words of the corpus in similar forms it was recorded into the corpus as the verb “let” and the pronoun “us”. The same applies to “don’t” which was recorded as the verb “do” and the adverb “not”. The verb “be” appeared also as “what’s” and “I’m” and both “what” and “is” and “I” and “am” were entered into the database as individual word types. Other verbs that appeared in the material were recorded into the corpus as they appeared. Table 6 presents the number of word families and total number of verbs in each of the five workbooks.

<b>Book</b>	<b>Word families</b>	<b>Total number of word tokens</b>
Come with me! 1	24	228
Come with me! 2	33	291
Go!	43	555
High five! 1–2	29	227
Jump in!	42	358

**Table 6.** Verbs.

*Come with me! 1* introduced the least verbs from different word families (24), of which three declinations (“am”, “are”, and “is”) belonged to the word family “be”. This meant that there were a total number of 26 different verb forms present in the material and a total of 228 appearances of verbs. This was only the second smallest number of total appearances. The verb “have” was the most common verb in the book with its 30 appearances, followed by “like” (25 times), “is” (23 times), “do” (21 times), and the rest of the verbs appearing 15 times or less. Six of the verbs in *Come with me! 1* were introduced only once. The verbs introduced in the workbook were all from the 1<sup>st</sup> frequency band.

The series’ second book *Come with me! 2* presented 33 word families of verbs and in addition to the declinations of “be” found also in *Come with me! 1* the head word “be” was also recorded three times. When each verb form was counted together there were 36 different verbs introduced in the workbook and a total of 291 appearances of verbs. “Can” was clearly the most common verb as it appeared 47 times, followed by “is” (39 times) and the other verb forms were introduced only 18 times or less. *Come with me! 2* introduced eight verbs that appeared only once and some verbs that were from higher frequency bands, but they will be discussed later in this thesis.

Like the other workbooks, *Go!* included several forms of the verb “be” and introduced a total of 49 verb forms and 43 word families resulting in 555 total appearances, which was 197 appearances more than the second most appearances in the material. Most of the verbs were presented in the base form and six verbs (“eats”, “loves”, “makes”, “plays”, “says”, and “washes”) in the third-person singular, out of which the words “eat” and “play” are also found in their base form. The word family “play” was the most common verb in the material with its 65 appearances. “Is” appeared 56 times, “do” appeared 20 times, and “can” appeared 36 times. The other verb forms were introduced 25 times or less. There were 11 verbs that appeared only once and five verbs that are not from the 1<sup>st</sup> frequency band.

*High five! 1–2* introduced 29 different word families and 31 different forms of verbs. *High five! 1–2* presented the fewest verbs in the workbooks for a total of 227 appearances, which was only one less than in *Come with me! 1*. The verb “be” appeared as “am”, “are”, and “is”, while the other verbs appeared only in one form. “Is” was found 29 times, being the most

common verb form in the material, followed by “go” with 28 appearances. Only four of the 31 verb forms in *High five! 1–2* were introduced once – “am”, “back”, “meet”, and “wearing” – the latter being the only verb to appear in the -ING form in *Go!*. “Do” appeared 19 times, “start” 11 times and both “play” and “sing” 10 times, and the rest of the verbs were presented nine times or less. Only one verb, “swing”, was from the 2<sup>nd</sup> frequency band, the rest being from the 1<sup>st</sup>.

There were 51 different verb forms found in *Jump in!*, but only 42 different word families. Just like in the four other workbooks, the verb “be” was encountered as “am”, “are”, and “is”, but also as “be”. The “have” word family included the base form “have” and the past simple “had”, and the “go” word family included the base form “go” and the third-person singular “goes”. Altogether there were 358 times when verbs were presented in *Jump in!*. “Start” was the most common verb, appearing 36 times. “Do” appeared 27 times, “like” appeared 22 times and “are” a total of 22 times. The rest of the appearances occurred fewer than 20 times. *Jump in!* included four verbs that were not from the 1<sup>st</sup> frequency band (“clap”, “skip”, “slap”, “stomp”).

#### 4.4 Adjectives

The five workbooks presented a total number of 75 different adjectives. The number of word tokens in the five different workbooks ranged from 122 times to a total of 371 encounters. The group of adjectives gathered into the corpus could be grouped into two sections – colours and opposites. ‘Colours’ presented variation as only 3/5 workbooks introduced “purple” and one left “grey” out. Altogether a total of 75 different adjectives were encountered in a total number of 1,170 appearances. “Blue” was the most common adjective in the material appearing 76 times, followed by “green” and “red” which both appeared 69 times in the workbook. 27 adjectives out of the 75 different word families appeared only once in the material meaning that only one workbook had chosen to include the adjective in their teaching material. Table 7 presents the number of word families and total number of nouns in each of the five workbooks.

Book	Word families	Total number of word tokens
Come with me! 1	24	122
Come with me! 2	34	131
Go!	43	330
High five! 1–2	35	216
Jump in!	29	371

**Table 7.** Adjectives.

*Come with me! 1* presented the fewest adjectives (24) and the least word tokens (122). Five of the adjectives introduced in the workbook were from the 2<sup>nd</sup> frequency band and the other 19 were from the 1<sup>st</sup> frequency band. Both “blue” and “purple” appeared 13 times in the material making them the most common adjectives. The appearances of different colours were not even as “brown” appeared only three times. Eight adjectives appeared only once in the material and five of those adjectives belonged to the 2<sup>nd</sup> frequency band. All the other adjectives in *Come with me! 1* were from the 1<sup>st</sup> frequency band.

A total of 29 different adjectives were encountered in *Jump in!* which was the second least, but the total number of appearances was the largest with 371 encounters. “Extra” was the most encountered adjective with its 37 appearances, followed by “happy” with its 32 appearances, followed by the colours “blue” and “red” with their 27 appearances respectively. Like *Come with me! 1*, colours were not presented in equal amounts as “white” appeared only 16 times and both “brown” and “grey” (and “gray”) were encountered a total of 17 times each. Eight adjectives appeared only once and five of these were found in the additional exercises at the end of the book. The adjectives in *Jump in!* belonged mainly to the 1<sup>st</sup> frequency band, with a few from the 2<sup>nd</sup> and 3<sup>rd</sup> frequency bands and one from the 7<sup>th</sup> (“undercover”).

*Come with me! 2* did not recycle colours nearly as much as the series’ first book did as “blue” was the most common colour in the workbook with its five appearances and both “black” and “red” appearing just once. There were 34 different word families and the total number of appearances within the book was 131. The most encountered adjectives in the workbook were the opposites “happy” and “sad” that were each found 15 times. The other adjectives

were found only 10 times or less, out of which 11 adjectives appeared only once. There were just two adjectives in *Come with me! 2* that were from the 2<sup>nd</sup> frequency band (“amazing”, “brave”), the others were from the 1<sup>st</sup>.

*High five! 1–2* introduced 35 different adjectives and a total number of 216 word tokens. “Ready” and “steady” were the most common adjectives, both appearing 13 times respectively, followed by “bad” with its 12 appearances and “orange” appearing 11 times. The other adjectives were each found 10 times or less. Six of the adjectives appeared only once and four adjectives were from the 2<sup>nd</sup> frequency band, the most common adjective “steady” being one of them. The other 29 adjectives were from the 1<sup>st</sup> frequency band.

*Go!* introduced the most different word families (43) and the workbook included the second most word tokens (330). *Go!* had the most variety in its adjectives from the perspective of frequency bands as there were adjectives from the 1<sup>st</sup>, 2<sup>nd</sup>, 4<sup>th</sup>, 6<sup>th</sup>, and 7<sup>th</sup> frequency bands, but whether they were appropriate adjectives for young learners to learn will be discussed in section 5 of this thesis. “Nice” was the most common adjective, appearing 23 times. Different colours were encountered 15 to 21 times, “blue” and “red” being the most common adjectives. “Big” was encountered 16 times and the rest of the adjectives 9 times or less each. 11 adjectives appeared only once in the workbook.

#### 4.5 Adverbs

There were 21 different words in the material that were categorised as adverbs. Like the pronouns of the vocabulary, all the adverbs in the corpus belong to the 1<sup>st</sup> frequency band. Only “all”, “here”, “not”, “out”, “there”, and “too” appeared in more than one workbook, the other 15 adverbs appeared in only one of the books. The total number of appearances of adverbs in the whole corpus was 130. “Not” was the only adverb featured in all five workbooks, appearing 63 times in total, which was most of individual adverbs. “Not” was combined with the verb “do” and was encountered in the form “don’t”. “Do” and “not” were categorised as separate words as they belong to different parts of speech and to maintain comparability within the corpus. As “not” appeared with the verb “do” in the material, “not”

was categorised as an adverb as it relates to a modified verb. Table 8 presents the number of word families and total number of adverbs in each of the five workbooks.

<b>Book</b>	<b>Word families</b>	<b>Total number of word tokens</b>
Come with me! 1	4	9
Come with me! 2	6	27
Go!	10	43
High five! 1–2	6	23
Jump in!	8	28

**Table 8.** Adverbs.

*Go!* introduced both the most adverb word families (10) and the most individual appearances of adverbs (43). Even though adverbs appeared in *Go!* for a total of 43 times, only four of those adverbs (“after”, “here”, “not”, and “too”) appeared more than once and only “not” and “too” appeared in more than one chapter of the workbook. *Jump in!* introduced only one word family less (8) than *Go!* but the total number of adverbs that appeared in the workbook was only 28.

*High five! 1–2* introduced six different word families but managed to introduce individual word tokens 23 times. Only two of the adverbs (“ahead” and “not”) in *High five! 1–2* appeared more than once in the material but in this case both adverbs appeared in two of the chapters rather than in just one section of the workbook. “Here”, “later”, “maybe”, and “once” each appeared only once, and the latter three being the only uses of those adverbs in the whole material.

*Come with me! 1* only introduced four word families of adverbs (“all”, “late”, “not”, and “now”) of which “not” (5 times) and “now” (2 times) appeared more than once in the workbook. *Come With Me! 2* included six different word families (“back”, “here”, “not”, “out”, “too”, and “well”), where “not” was introduced 21 times, “here” twice, and the other four only once. If the two *Come with me!* workbooks were analysed together their total number of word families would be ten, which would be more similar to the other books that have combined both 1<sup>st</sup> and 2<sup>nd</sup> grade material together unlike the *Come with me!* series.

The second most common adverb in the material was “ahead” which was found 13 times only in *High five! 1–2* as it appears in the title of the chapter (“Go ahead!”) preceding the first official chapter of the workbook. The adverb that appeared the third most was “too” with its 12 appearances in two different books. “Too” was the second most recycled adverb as it appeared in nine different chapters in *Go!* but at the same time only once in *Come with me! 2*. The rest of the adverbs appeared less than 10 times each with over half of them being introduced only once.

#### 4.6 Prepositions

The five workbooks introduced 14 different prepositions that appeared altogether 117 times. All the words categorised as pronouns appear in the 1<sup>st</sup> frequency band. “To” appeared in each of the five books being the only preposition to do so. “On” appeared in four of the five materials and both “in” and “with” were found in three of the workbooks. Table 9 presents the number of word families and total number of prepositions in each of the five workbooks.

Book	Word families	Total number of word tokens
Come with me! 1	3	4
Come with me! 2	6	13
Go!	7	55
High five! 1–2	5	19
Jump in!	8	26

**Table 9.** Prepositions.

“On” was the most common preposition in *Jump in!* appearing a total of seven times while “with” only appeared once in the workbook. The other six prepositions in *Jump in!* appeared two, three, or four times with the total number of appearances in the book resulting to 26. Different books preferred different pronouns as “on” appeared only once in both *High five! 1–2* and *Go!*, and only twice in *Come with me! 2*. Half of the prepositions in *Jump in!* appeared in only one chapter while the other half each appeared in two chapters of the workbook.

*Come with me! 1* introduced only three different prepositions with a total number of four word tokens of any prepositions in the whole workbook, which was clearly the least of the



five workbooks. “To” and “in” appeared only once and “with” twice, which excludes the “with” in the title of the workbook. *Go!* presented seven different prepositions that appear altogether 55 times in the material. The preposition “to” appeared a total of 33 times, being the most common preposition in any of the five workbooks. The other prepositions in *Go!* appeared significantly less with “at” appearing nine times, “for” eight times, “in” twice and the other three prepositions once each.

*High five! 1–2* incorporated 19 prepositions, ten of which were “at” which appeared in four different chapters of the book. “In” appeared five times in *High five! 1–2*, “to” twice and both “on” and “upon” only once respectively. *Come with me! 2* included six different prepositions with a total number of 13 word tokens. “With” appeared the most with four occurrences, “to” followed with three representations and the other four prepositions appeared either once or twice.

#### 4.7 Conjunctions

The part of speech that clearly introduced the least separate word families was conjunctions. The total number of word tokens was 157, which was the third smallest of the different parts of speech presented in the material. There were differences in the number of different conjunctions introduced ranging from one to four but also in how many times conjunctions appeared in each of the five workbooks. “And” was the only conjunction that appeared in each book with four appearances in one of the books and 55 in another. Each conjunction introduced belongs to the 1<sup>st</sup> frequency band. Table 10 presents the number of word families and total number of conjunctions in each of the five workbooks.

Book	Word families	Total number of word tokens
Come with me! 1	2	5
Come with me! 2	1	7
Go!	4	62
High five! 1–2	1	14
Jump in!	2	69

**Table 10.** Conjunctions.

*Come with me! 1* introduced “and” only four times and each of these times the conjunction was used in a short listening task where the dialogue could be found in the workbook. Each of the four times was however in a different chapter. The conjunction “or” was introduced only once in the phrase “cheese or chips” in an exercise where the letter combination “ch” was practised. These appearances make up a total of only five conjunctions used in the whole *Come with me! 1* workbook. *Come with me! 2* only introduced “and” a total of seven times in six different chapters of the workbook. The conjunction was used in examples on how to do a speaking exercise (“I like yellow and orange”) and like *Come with me! 1* in the dialogue of listening tasks and when practising a letter or sound.

*High five! 1–2* also only introduced “and” which appeared 14 times in five different chapters. *Jump in!* and *Go!* stand out from the other three workbooks as conjunctions appeared 69 times in *Jump in!* and 62 times in *Go!*. “And” appeared 53 times in *Jump in!* between items in a list and in song lyrics such as “And a moo, moo there,” and “If you’re happy and you know it,” which happened to also be the only place in the workbook where “if” was introduced.

“And”, “but”, “if”, and “or” appeared in *Go!* with “and” being the most common conjunction as it appeared 55 times in total, but oddly only in the first and last chapter of the workbook. “If” and “or” appeared three times each and “but” only once. *Go!* introduced “and” in song lyrics and lists while “if” and “or” both appeared in the same lyrics in only one of the chapters. “But” appeared in the example of a dialogue exercise in the phrase “But I can dance”.

#### 4.8 Interjections

The teaching material introduced 331 word tokens and a total of 17 different interjections out of which two (“shh” and “wow”) were not found in the BNC/COCA word family lists but were categorised as interjections, nonetheless. The number of different interjections varied from seven to eleven and their appearances throughout the workbooks ranged from 45 to 108. The interjections recorded into the corpus included a few words (“hooray”, “welcome”, “yuck”, and “yum”) which do not belong to the 1<sup>st</sup> frequency band like the other interjections that were encountered. Each book included the word forms “hello” and “hi” which both belong to the same word family “hello”. The other books apart from *Jump in!* included both

“thanks” and “thank (you)” and these words also belong to the same word family “thank”. Table 11 presents the number of word families and total number of interjections in each of the five workbooks.

Book	Word families	Total number of word tokens
Come with me! 1	10	48
Come with me! 2	11	76
Go!	7	108
High five! 1–2	10	45
Jump in!	6	54

**Table 11.** Interjections.

*Jump in!* presented six different word forms, which was the least of the five workbooks. There were seven different words, as there were two different forms that belonged to the word family “hello”. The word “bye” appeared three times as “bye-bye” and was counted as six different appearances. No other words were hyphenated so the words in the corpus are comparable to one another because of this decision. “No” appeared 14 times, which was the most of the interjections in *Jump in!*. “Yes” appeared 13 times and “hello” 11 times, and the other interjections six times or less. Six of the words were from the 1<sup>st</sup> frequency band but “hooray” is only from the 9<sup>th</sup> frequency band.

There were seven word families introduced in *Go!* and nine different word types in total. The most common interjections were the opposites “yes” and “no” with a total of 23 appearances each, followed by “bye” with 22 appearances. Only “welcome” belonged to the 2<sup>nd</sup> frequency band while the other interjections are from the 1<sup>st</sup> frequency band. While “welcome” appeared seven times, which was not the least, it appeared in only the second chapter of the workbook. “Please” was the least common interjection and appeared three times, but in three different chapters. Even though there was only one word family more than in *Jump in!* the total number of appearances in *Go!* was double the amount for a total of 108 appearances.

*High five! 1–2* introduced ten different word families, two of which included two different words. There were only 45 appearances of interjections in total, which was the least of the five workbooks. Seven of the 12 word types appeared only in one chapter and five of those seven words were introduced only once. Five of the 12 words were used once, one of which was from the 9<sup>th</sup> frequency band (“yum”) and one that was from the 10<sup>th</sup> (“yuck”). “Yes” was the only interjection to appear more than ten times as it was encountered 13 times. “No” appeared eight times, “please” seven times and the rest either four times or just once.

*Come with me! 1* had ten different word families and 12 different words that appeared 48 times in total. *Come with me! 2* included 11 word families and 13 different words and they appeared 76 times in total. The series’ two books had almost the same interjections appearing in them, but there were few words that appeared in only either one of the books. “Cool” appeared only in *Come with me! 1* and “sorry” and “shh” appeared in *Come with me! 2*. “Yes” was the most common interjection in both two workbooks appearing 17 times in the 2<sup>nd</sup> grade workbook and 12 times in the series’ first workbook. “No” was the second most common interjection in both books as well. Four of the interjections in *Come with me! 1* and three interjections in *Come with me! 2* appeared only once. Both workbooks included one interjection that belonged to the 2<sup>nd</sup> frequency band, “welcome”. “Wow” and “oh”, both of which appeared in both workbooks, were not found in the BNC/COCA word lists, and thus cannot be categorised into any frequency band. The interjection “shh” found in *Come with me! 2* also is not found in any of the BNC/COCA words lists and is not linked to any frequency band.

#### 4.9 Determiners

The determiners that appeared in the teaching material were categorised into three main sections and one section tagged as ‘other’, with words that could not be set under the other three headings that were ‘articles’, ‘numbers’, and ‘possessive determiners’. *Come with me! 2* presented 26 different word families which was the most of the five workbooks, but the total number of word tokens of determiners was the least (221). *Come with me! 1* had 16 different word families, as did *Go!* and *High five! 1–2*. *Jump in!* had a few more with its 18 different word families of determiners. The total number of determiners that appeared in the

material was 1,811 and they appeared thoroughly throughout each chapter of each workbook. All the determiners are found in the first frequency band. Table 12 presents the number of word families and total number of determiners in each of the five workbooks.

Book	Word families	Total number of word tokens
Come with me! 1	16	282
Come with me! 2	26	221
Go!	16	422
High five! 1–2	16	346
Jump in!	18	540

**Table 12.** Determiners.

Articles appeared in the teaching material 1,074 times. “The” was clearly the less common article with only 106 appearances which amounts to approximately 9,87% of the use of articles in the five workbooks. “A” appeared 968 times amounting to the largest portion (90,13%) of the use of articles. Both “a” and “the” are found in the 1<sup>st</sup> frequency band. The use of “an” was scarce and is grouped under the head word “a” as they belong to the same word family.

*Jump in!* presented some of the vocabulary with the indefinite article (275 times) such as “a bed”, “an ear”, and “a green pencil” but at other times there was no article for example with “apples” or “games”. The definite article was found 28 times for example in songs, and an activity called “slap the card”. *Go!* opted to teach vocabulary by presenting words with either the indefinite article “a” or emitting the article altogether with words that do not require it such as “toast” or “shoes”. “A” appeared 219 times and “the” only 38 times, which was still the most of any of the books. “The” was used with hobbies such as “the piano” and “the drums”, and in songs.

*High five! 1–2* had only seven occurrences of “the” which were appeared in lexicon-grammatical units such as “in the morning” and “knocking on the door” which was found in a song. The word family “a” was presented 218 times out of which there were seven uses of “an” with the nouns “apple”, “elephant”, and “orange”. *High five! 1–2* introduced the

majority of the workbook's nouns with the indefinite article "a" such as "a brother", "a school bag", "a sofa", and "a cookie". *Come with me! 1* chose to present food items in the plural form but other nouns such as "a dog" and "a leaf" included the indefinite article "a". The workbook introduced the bugs that appear as the main characters of the series with the definite article, for example "the ladybug" and "the spider".

Numbers were categorised as determiners as they appear mainly before a noun determining the quantity of the noun. At other times numbers appeared individually as "three" or "ten" for example in the early chapters of the workbooks when numbers were encountered for the first time. Of the numbers that appeared in each of the workbooks "two" was the most common number in the whole material appearing 72 times while "nine" appeared the least (30 times). The smaller the number was the more it appeared, and any number above 7 appeared less than 35 times throughout the whole material. Like the articles "a" and "the", numbers appeared evenly throughout the workbooks with some numbers appearing more frequently than others.

Each book included the numbers one through ten. *High five! 1–2* also included "eleven" and "twelve", *Come with me! 2* introduced numbers up to "twenty". Numbers were only recorded if they were spelled out (e.g., "four" or "nineteen") and digits (e.g., 4 or 19) were excluded from the corpus. "One" was the most recycled number only in *Jump in!* with its 31 appearances. *High five! 1–2* recorded 12 uses of "three" and *Come with me! 1* introduced it 13 times. "Two" was the most common number in *Go!* appearing 14 times and "twenty" in *Come with me! 2* appearing 12 times.

The last clear group of determiners, possessive determiners, were categorised as determiners rather than pronouns because of their similar function to numbers. "My", "their", and "your" appeared before nouns to determine the possession of the item or other matter that the noun was referring to. "My" and "your" appeared in each book with "my" being the more common determiner with its 89 appearances and "your" appearing less than half of that with its 43 uses. "Their" appeared only once in *Come with me! 2* in the phrase "Where's their mother?" being the only use of it in the material in a section where the letter combination "th" was practised.

The other words grouped as determiners were: “all”, “many”, “more”, “much”, “some”, and “what”. Two uses of “what” were categorised as a determiner as they were used in the lexical units “What a noise!” and “What a big spider!”. “Many” appeared in three of the materials for a total of 24 times which was most of the uncategorised determiners. “All” appeared six times but only in *Jump in!* and the rest of the words only once or twice in only one of the five workbooks.

## 5. Discussion

The results and the corpus help answer the three research questions. The first research question is supported and answered by its sub questions that focus on the details of the vocabulary in the material. The second research question focuses on frequency bands and the third research question on vocabulary recycling.

### 5.1 Vocabulary in the Early Language Learning Material

The first research question is “What vocabulary is introduced in the early language learning material?”. The sub questions “How many word tokens are introduced?”, “What parts of speech are covered?”, and “To which topics does the vocabulary belong?” support the first main research question.

The material, which consists of five workbooks for young learners, includes a total number of 9,377 word tokens that are referred to as words in this thesis. The number of word tokens in the material expresses the total word count with every repetition recorded (Brysbaert et al., 2016, p. 1). The total number of tokens consists of the different word types presented in the five different workbooks. *Go!* included the most word tokens of the five workbooks with the total number of 2,654 word tokens which is 28,30% of the total number of word tokens in all of the material. The number of word tokens is over 1/4 of all the material’s tokens and not expected as there are five different workbooks. It would be expected *High five! 1–2*, *Go!*, and *Jump in!* would have the most word tokens as they are to be used in both 1<sup>st</sup> and 2<sup>nd</sup> grade and it would also provide information that young learners learn the same number of words no matter which workbook is chosen. The *Come with me!* series should have less word tokens than the other three books as each of the two books is designed to be used in one grade rather than two.

*Jump in!* follows second with 2,493 tokens (26,59%) and *Come with me! 2* includes the third most (1,561) word tokens (16,65%). There is approximately a 10% gap between the percentages of the second and third most word tokens. As the two workbooks have the same number of pages (96 pages respectively) in them, the difference between the number of word tokens they introduce is too large. *High five! 1–2* has the second fewest (1,427) word tokens



(15,22%) and *Come with me! 1* introduces 1,242 word tokens which is the fewest of the material and only 13,24% of the total number of word tokens of the material. *Come with me! 1* includes the fewest number of word tokens but the workbook also has the fewest pages (88) of the five workbooks, which is one reason why it introduced the smallest number of word tokens. However, *Come with me! 1* has only eight pages less than the other workbooks and therefore this cannot be only reason for providing less word tokens.

Unlike *Go!*, *High five! 1–2*, and *Jump in!*, the *Come with me!* series offers separate books for 1<sup>st</sup> and 2<sup>nd</sup> grade. As *Come with me! 1* is only designed for the 1<sup>st</sup> grade, it is not surprising that it includes fewer word tokens than the workbooks designed to be used by young learners in both 1<sup>st</sup> and 2<sup>nd</sup> grade or *Come with me! 2* used in the 2<sup>nd</sup> grade. If the word tokens in both *Come with me! 1* and *Come with me! 2* are added together for a total of 2,803 word tokens, the number would be the largest number of word tokens in the material (29,89%), but not much larger than what *Go!* offers. As discussed above, also *Come with me! 2* is designed to be used in only one grade (2<sup>nd</sup> grade) and thus the number of word tokens (1,581) is justifiably to be smaller than those books that are aimed to be used in both the 1<sup>st</sup> and 2<sup>nd</sup> grade. However, *High five! 1–2* introduces only 1,427 word tokens which is less than *Come with me! 2*.

The chapters in *High five! 1–2* focus more on the ten core words taught in each main chapter rather than including longer lexical units or more complex sentences in English. The choice to focus on the ten core words may be the result of understanding that young learners are still learning to read in their first language when starting their foreign language learning journey and should increase the number of words they learn instead of other aspects of grammar. However, the workbook includes a lot of pictures which may relate to a lot of discussion based on the pictures. The tasks in *High five! 1–2* focus on listening, drawing, and learning vocabulary through games, but not writing. Students are not writing down the words of the chapters in tasks or any other way in the workbook and are not encountering the words through this type of recycling.

Learning a foreign language revolves around learning enough vocabulary (Norberg & Nordlund, 2018, p. 463). While on one hand the number of word tokens in each workbook

suggests that there are plenty of opportunities to encounter English words in the material, the number does not provide any factual evidence on how many word tokens are realistically encountered when using the material. On the other hand, it cannot be expected that young learners read and encounter each word token in the material as teachers are likely to select the best exercises from the workbooks rather than having the students work on each task. The number of word tokens should be compared to more foreign language learning material, that is aimed especially at young learners, to understand better whether the number of tokens is low or high, or even enough to learn vocabulary. Nation and Anthony recognise that when young learners learn even a few new words, the impact is more considerable as their total vocabulary size is so small (2017). Therefore, while some of the material offers more word tokens to encounter, encountering any word types that the material introduces is beneficial for young learners.

Each workbook had words that were unique to that specific teaching material and did not appear in any of the other workbooks. The goal of language learning is to develop proficiency (Nikolov & Djugunović, 2019) and to meet the goal various parts of speech and words must be encountered. The parts of speech covered by the thesis material are nouns, pronouns, verbs, adjectives, adverbs, prepositions, conjunctions, interjections, and determiners.

There were 3,230 word tokens that were nouns, making nouns the most common part of speech in the material. This amounted to 34,45% of the whole material's word tokens, which is just over 1/3 of the total number of tokens in the material. The themes and topics the vocabulary belongs to is discussed under the next sub question. Nouns being the most frequent word type in the material is not a surprise as nouns are the most common part of speech in the English language (Hudson, 1994). Some of the nouns that appeared in only one of the workbooks are more surprising (e.g., "farm", "homework", and "strawberry") than others (e.g., "oink", "hopscotch", and "bingo") because the surprising word types closely relate to the topics of the material.

Pronouns were the fifth largest part of speech in the material with 772 word tokens. Pronouns are essential to grasp when young learners learn to talk about themselves and other people in their lives. Pronouns make up only 8,23% of the materials word tokens. 42,10% of the word

tokens that were pronouns are made up of the word family “I”. Pronouns were not a part of the core words taught in each main chapter, which makes it understandable that pronouns do not appear as much as nouns, verbs, adjectives, and certain determiners in the material.

Verbs were the third largest group of part of speech in the material behind determiners. Verbs are essential to learn if young learners want to form sentences that describe what they are doing or even to ask the question “How are you?” that appeared often in the material. 1,659 verb word tokens amount to 17,69% of the word tokens of the whole material. The verb “be” appeared in its various declinations which made it the most common verb in the material. “Be” being the most common verb signals that young learners should know how to express themselves and describe others and the world surrounding them.

Colours were the most frequently appearing adjectives along with very general descriptions such as “nice”, “old”, and “fun”. The total number of word tokens that were adjectives was 1,170 which is the fourth most. Colours was the topic that appeared very early on in each of the workbook. *Go!* introduced ice cream flavours such as “chocolate”, “lemon”, and “pear”, which were unique to that workbook, just like every workbook had adjectives that only appeared in one of the workbooks.

Adverbs, prepositions, conjunctions, and interjections make up the minority of word tokens in the material. Altogether they make up 7,84% of all the word tokens in the material. The adverb “not” appeared in phrases such as “I don’t like football”, making it appear as many times as it did in the material. Other than that, adverbs were not treated as being relevant vocabulary for young learners to learn. Interjections, such as “hello”, were the only part of speech of these four to appear in the core vocabulary of introductory chapters. The other parts of speech were introduced in songs, speech exercises – especially “yes” and “no” – and short dialogues or comics. The way these parts of speech appear in the text is a clear connection to why they make up such a small part of the total number of word tokens. When comparing the number of word types and their appearances in each of the material, more reasonably conjunctions such as “and” appear in every workbook than the noun “map” because there is much less variety in the total number of word families that are conjunctions.

Determiners are the second most common parts of speech type in this study, which is mainly due to categorising the articles “a”, “an”, and “the” as determiners. These articles appear with the core vocabulary (e.g., “an ear”, “a frog”) taught in each of the workbooks and thus it results in large numbers of each article appearing in the material. *Come with me! 2* was the only workbook in which the articles amounted to less than half (42,99%) of the total number of tokens that were determiners. This is because *Come with me! 2* introduced the most numbers (1–20) of the workbooks and they overtook articles in the total number of word tokens. Numbers and possessive determiners make up most of the rest of the determiners, which only leaves occasional determiners such as “many” and “more” to appear only a few times. Numbers appeared in various tasks to determine the number of objects, and possessive determiners in songs and phrases such as “What’s your name?”.

The vocabulary taught in each workbook was categorised into different topics, if there was a clear theme seen. Barclay and Schmitt consider teaching situational vocabulary necessary (2019) and in this thesis the workbooks’ topics define the contexts where vocabulary is used. This type of categorisation only applies to nouns, a selection of adjectives, and numbers that are categorised as determiners in the database. Demircioğlu argues that vocabulary should be met in appropriate contexts for young learners to learn the words and in order to build connections between words and where to use them accordingly (2010, p. 440). The topics encountered in the material mainly revolve around the assumed daily life of young learners, but vocabulary relating to modes of “moving around” (e.g., “boat”, “skateboard”, “tractor”) was the most farfetched topic, but at the same time not entirely irrelevant. Like Norberg and Nordlund in their 2018 study of English learning material in Sweden, the material in Finland also revolves around the assumption that young learners need to know vocabulary relating to topics such as animals, colours, and family.

An often occurring “organising principle” with young learners is *topics* (Rixon, 2019, p. 284). Each of the five analysed workbooks had various, meaningful topics to which young learners can easily associate. These topics ranged from people and animals with vocabulary about family, friends, and pets, to items such as clothes and food, and vocabulary relating to school and classrooms. Additionally, different word category items such as colours, a moderate number of verbs and adjectives, numbers one through twenty, and weekdays were

introduced as separate topics as well. Rixon has noticed that topics are not as common with older learners (2019), which can mean that older learners start focusing more on linguistic characteristics as they have gathered enough vocabulary knowledge through different topics.

The use of topics with young learners can be justified as not only being relevant to the lives of young learners but topics are an efficient way of introducing certain lexicon-grammatical units (Rixon, 2019, p. 284). *Jump in!* introduced verbs as a topic with the title “I can” and *Go!* introduced verbs that relate to activities you can do in the school playground. Verbs were thus connected in a meaningful way to the lives of young learners rather than just listing verbs that must be learned. Demircioğlu highlights the importance of providing enthusiastic ways of learning vocabulary to engage and motivate students to learn a foreign language (2010). Learning vocabulary that is relevant to the young learners’ lives is more motivating than learning words that they cannot find a connection to.

The number of different ‘animals or pets’ mentioned in the material differed with each book, but all the workbooks included the topic as one of their main chapters. *Come With Me! 2* introduced the number least animals (“cat”, “pig”, & “rabbit”), but it may rely on students remembering them from the series 1<sup>st</sup> book which introduces 18 animals. Young learners must be expected to be animal lovers as each of the workbooks offer vocabulary on the topic.

*Come with me! 2* and *Go!* do not introduce any topic related to ‘body parts’. The other three workbooks introduce words such as “an ear” and “a head” in their own chapters. Just like in the topic above, the authors of *Come with me! 2* expect the young learners using the book to be familiar with the series’ first workbook and thus with the vocabulary introduced in it. ‘Body parts’ are not a part of the vocabulary in *Go!* for an undisclosed reason while the vocabulary belonging to the topic in the three workbooks is fully from the 1<sup>st</sup> frequency band. As young learners learn to talk about themselves it is natural to learn vocabulary such as “an eye” and “legs” relating to their bodies.

All five workbooks introduce ‘clothes’, but once again *Come with me! 2* has only a few new words it introduces, relying on the fact *Come with me! 1* has already introduced the main vocabulary. Nonetheless, words should not only be recycled within a workbook but also

throughout the series to guarantee learning. Teaching clothes-related vocabulary may be common in the teaching material as colours are easily linked to discussion of clothes. Basic colours are expected to be taught early on for young learners, which is also the case in this material.

'Colours' is one of the topics that has its own chapter in each of the books of the two topics that are not nouns. There were eleven colours that were presented in the workbooks but *Come with me! 1*, *Go!*, and *Jump in!* only featured ten of the colours. *Go!* and *Jump in!* are missing "purple", and *Come with me! 1* is missing "grey". Some colours are unexpectedly used more often in the material than others. "Blue" was the most repeated colour in each of the books, covering 13,65% of the colours in the whole material. Some of the colours were prioritised differently within the material, with some colours being repeated throughout material for a total of eight times, while some only appeared three times.

'Family and friends' is one of many themes represented in *Come with me! 2*, *High five! 1–2*, and *Jump in!*, but surprisingly the theme is not present in *Go!* or *Come With Me! 1*. Teaching young learners should revolve around topics that relate to their lives and the vocabulary of the 'family and friends' theme is one that could easily be justified to be taught. Young learners learn to talk about themselves and thus learning family words is important. The word "dad" appears twice in *Go!* in a picture where a man has a shirt with the text "World's greatest dad", but other than that no mention of any family members is seen in the material. Of course, *Come with me! 2* resolves the issue of not introducing family related vocabulary in the series' first book, but as the theme is found in the first half of *High five! 1–2* and *Jump in!* it would suggest that the topic is relevant already in the 1<sup>st</sup> grade.

'Food'-related vocabulary can be found in each of the workbooks. Some of the words appeared as singular items ("an apple" in *High five! 1–2*) and plurals ("apples" in for example *Jump in!*). The topic of 'food' includes common breakfast foods and snacks, but the words taught in different books also varies. There are also different synonyms used for the same item in different workbooks, such as "a cookie" and "a biscuit".

*Come with me! 1* did not introduce any 'hobbies', which is a topic in the other four workbooks. Some of the topics are introduced only in one of the workbooks of the *Come with me!* series. The vocabulary consists of musical hobbies, sports, and different types of art forms. Some words (e.g., "ice hockey", "floorball") are undoubtedly introduced in the material because the workbooks are used in Finland, where the sports are very common.

Words that are items found in a 'house' or at 'home' were only introduced as a main chapter in *Jump in!*. The topic revolves around household items and furniture that are taught in the chapter, but the theme includes more specific words such as the proper nouns "an Xbox" and "an iPad" which are found in other sections of *Jump in!*. *High five!* also introduces furniture items. While words to describe items at home may be necessary, it is unlikely that young learners need to discuss about "armchairs" in English which could be why the topic cannot be found in other books.

Words relating to 'nature and outside' were not common in the books and only *Come with me! 1* had a whole chapter devoted to the topic as the main vocabulary. Young learners are still at an age when playing and spending time outside is expected of them, even if it only happens during recess time. Some of the words, such as "a tree" and "a rock", introduced in *Come with me! 1* luckily appear in the other workbooks, but not all word families and not in every workbook.

*Come with me! 1*, *Jump in!* and *Go!* introduce numbers from 1 to 10, but *High five! 1-2* has chosen to also include "eleven" and "twelve" in its book. This choice to include 11 and 12 is most likely because the workbook has 12 different chapters in it. The same reasoning does not apply to *Come with me! 2* which introduces numbers 1 to 20. As the workbook is designed for 2<sup>nd</sup> grade, it is a wise choice to introduce some more vocabulary related to numbers than to just reintroduce and recycle what is learned in the series' first book.

'School'-related words are not introduced in *Come with me! 1*, but the other four books include vocabulary of items found in the classroom. Barclay and Schmitt concluded that using vocabulary that is useful for young learners, in this case words relating to their classroom, helps students navigate in that setting (2019). This justification makes it unclear why such

vocabulary is not taught in *Come with me! 1*, but only in the series' second book. Young learners spend hours every day and each week at school and learning such meaningful vocabulary should be apparent.

The topic of 'transportation' explicitly appeared only in *Jump in!*, but the word "car" for example was found in all workbooks except *Come with me! 1*. Each time "car" was encountered it did not relate to the vehicle driven outside but it also appeared with the topic of 'toys'. Just as it was assumed that young learners still spend many hours outside, it can also be expected that these students spend time on their bicycles and kickboards, which justifies the need and want to learn such vocabulary while some of the words such as "a motorbike" does not relate to all the young learners' lives. The last clear topic that is introduced in *Come with me! 2*, *High five! 1–2*, and *Go!* is 'toys'. Young learners are still interested in toys which is a clear indication to teach such vocabulary to them.

While 'head names', 'items', 'jobs', 'miscellaneous', 'onomatopoeia', 'places', 'proper nouns', 'seasons', 'time', and 'weekdays' do not have individual chapters devoted to them, nouns were categorised under these topics. As the number of chapters in each workbook is limited, these topics have not been seen important enough to have a main chapter devoted to them. Instead, words related to these topics appear within chapters in songs, texts, and other exercises. The vocabulary introduced ranges from words that are very significant for young learners to words that are not that necessary for them to learn, but a very wide range of vocabulary is introduced in the five workbooks.

## 5.2 Frequency Bands of the Vocabulary in the Early Language Learning Material

The second research question is "Which word frequency bands do the words belong to?".

Understanding frequencies is necessary when deciding what to teach to language learners according to Nation (2021). Some words are more useful than others and this can be determined by word frequencies (Nation & Wearing, 1997). When scanning through the corpus, some of the vocabulary presented is noticeably chosen with other criteria than choosing words from appropriate word frequency bands for young learners. Barclay and



Schmitt suggest that how often words occur in the English language is what should guide the choice of the words that should be learned first by learners (2019, p. 802). This supports the claim that young learners should learn the most frequent words, which appear as the 1st and 2nd frequency band word families in the BNC/COCA word family lists. As the most frequent words appear the most, learning them provides students with the best text coverage and learning low-frequency words is less useful (Barclay & Schmitt, 2019, p. 802).

The material consists of 88 noun word families that are outside of the 1<sup>st</sup> and 2<sup>nd</sup> frequency bands such as “hopscotch” (15<sup>th</sup>), “oysters” (6<sup>th</sup>), “smoothie” (13<sup>th</sup>), and “panda” (9<sup>th</sup>). This leaves 265 word families that belong to the most frequent word families of the English language according to the BNC/COCA word family lists. Of course, teaching material that would consist only of the most frequent words would offer young learners the most optimal vocabulary to learn but understanding that nouns are the most common part of speech provides reasons as to why there are also words from other frequency bands. As 75,07% of the nouns are from the 1<sup>st</sup> and 2<sup>nd</sup> frequency bands, over 3/4 of the materials’ nouns are words that are especially relevant for young learners to learn. Proper nouns are not a part of the BNC/COCA family words lists and Brysbaert et al. wonder whether people recognise names and to what extent (2016, p. 10). There were ten proper nouns which are not from any frequency band, and it is hard to estimate their difficulty. As there are more nouns than any other words in the material there is more variety in the word frequencies of nouns. There are some words in the material that are clearly in the latter 1,000-word families but some of the words have for example a cultural reasoning for appearing in the books. In *Jump in!* the word “reindeer” is introduced, which belongs to the 9<sup>th</sup> frequency band, but reindeer are common animals found in Finland, which makes including the word more justifiable.

There is variety in the use of English within the five workbooks especially with nouns. *Go!* relies on the use of the generalised dialect of British English which can be concluded from the words choices made in the book such as “biscuit” rather than “cookies” and “trousers” instead of “pants”. The other four books tend to lean onto the generalised dialect of American English, but if individual lexical items are looked at it cannot be stated that each book uses vocabulary that relates to only one dialect as there is overlap between dialects. The choice of dialects results in different words that at times end up being from different frequency bands.

Some synonyms have more of a gap between their frequencies in the BNC/COCA word family lists, as “biscuit” is found in the 2<sup>nd</sup> frequency band while its synonym “cookie” is only in the 7<sup>th</sup> frequency band. “Rabbit” is a head word from the 1<sup>st</sup> frequency band while “bunny” is from the 6<sup>th</sup> frequency band. Some synonyms are found from consecutive frequency bands, “chips” is from the 1<sup>st</sup> frequency band and “crisps” is from the 2<sup>nd</sup> frequency band. “Rubber” is found in the 4<sup>th</sup> frequency band while its synonym “eraser” is found in the 5<sup>th</sup> frequency band. While neither of the words are in the scope of the relevant first words that should be taught to young learners, “rubber” is better because of its higher frequency.

Most of the vocabulary related to clothes is found from the 1<sup>st</sup> and 2<sup>nd</sup> frequency bands, but there are terms from the latter frequencies of word families. “Shorts” appear in three of the books and “pants” appear in two books, and these words are surprisingly found in the 5<sup>th</sup> frequency band. *Go!* has chosen “trousers” instead of “pants” and the term is found from the 4<sup>th</sup> frequency band and thus is found more frequently in texts overall, and surprisingly the word “jeans” used in *Jump in!* is found in the 2<sup>nd</sup> frequency band. *High five! 1–2* and *Come with me! 1* have chosen to include the word “hoodie” in their chapters about clothes and the term is only found in the 18<sup>th</sup> frequency band. If Nation’s rule of learning 1,000 word families annually, which is based on native speakers of English (2001), young learners would be in their 20s before needing to learn word families from the 18<sup>th</sup> frequency band.

From the words categorized to be ‘family’-related, all words except one were from the 1<sup>st</sup> frequency band. *High five! 1–2* decided to use the word “granny” for “grandmother” and had thus chosen a word from the 6<sup>th</sup> frequency band. If “granny” was spelled as “grannie” it would have been from the 1<sup>st</sup> frequency band as the head word “grandmother” is from the 1<sup>st</sup> frequency band. *Come with me! 2* was the only other book to include any word for “grandmother” and opted to use “grandma”, which is found from the 1<sup>st</sup> frequency band like the other words from the ‘family’ topic.

‘Nature’ vocabulary mainly includes words from the 1<sup>st</sup> and 2<sup>nd</sup> frequency bands, but such a common feature that children draw and is present in clothing, “rainbow”, is from the 5<sup>th</sup> frequency band as is “leaf”. The characters in the *Come with me* series are insects that are mentioned only once in the beginning on *Come with me! 1*. A few of the characters are from

the latter word family lists as “bumblebee” is found in the 14<sup>th</sup> frequency band and “ladybug” from the 18<sup>th</sup> frequency band – the synonym “ladybird” is found in the 15<sup>th</sup> frequency band. Norberg and Nordlund note that learning material for young learners may be constructed with the idea of including the language these learners use rather than focusing on the appropriate frequencies of words (2018, p. 469).

Most of the hobby-related words are found in either the 1<sup>st</sup> or 2<sup>nd</sup> frequency bands, but especially some sports are found in 3<sup>rd</sup>, 4<sup>th</sup>, or 5<sup>th</sup> frequency bands. *Go!* introduces “badminton” which is found in the 10<sup>th</sup> frequency band and “floorball” is not found in the first 20 frequency bands. These two sports are however popular in Finnish primary schools and especially floorball is a common hobby for Finnish people, which gives more incentive to teach the word while it is not a high-frequency word in English.

Most verbs are from the 1<sup>st</sup> frequency band. “Ski” and “skiing” which are from the same word family are from the 2<sup>nd</sup> frequency band and are still high-frequency words. Because of the snow and long winters in Finland, “ski” would have most likely been introduced in the material no matter the frequency band as it is a very popular sport amongst all ages. Five other verbs, “mix”, “slide”, “scream”, “seek”, and “swing” are also from the 2<sup>nd</sup> frequency band making the total number of word families from the 2<sup>nd</sup> frequency band six.

Only five verbs are not from the 1<sup>st</sup> and 2<sup>nd</sup> frequency bands. These verbs are “skip” (4<sup>th</sup>), “slap” (4<sup>th</sup>), “clap” (5<sup>th</sup>), “skate” (5<sup>th</sup>), and “stomp” (8<sup>th</sup>). These verbs are not used every day and it is not a surprise that they are not high-frequency words, but it is surprising that such words are taught to young learners. “Slap” was introduced in the name of game which most likely includes slapping and while the word is from the 4<sup>th</sup> frequency band young learners are likely to learn it as it is connected to the action it describes.

Like verbs, most of the adjectives in the material are from the 1<sup>st</sup> frequency band. 12 of the 75 different word families of adjectives are from the 2<sup>nd</sup> frequency band. “Pink” is the most common adjective from 2<sup>nd</sup> frequency band and the only colour not from the 1<sup>st</sup> frequency band. The seven adjectives not from the 1<sup>st</sup> or 2<sup>nd</sup> frequency band are “mobile” (3<sup>rd</sup>), “lemon” (4<sup>th</sup>), “mini” (6<sup>th</sup>), “pear” (6<sup>th</sup>), “raspberry” (7<sup>th</sup>), “undercover” (7<sup>th</sup>), and “vanilla” (7<sup>th</sup>).

“Lemon”, “pear”, “raspberry”, and “vanilla” were introduced as ice cream flavours and thus appear as adjectives. The frequency bands they are found from are surprising as they are common ice cream flavours, at least in Finland, and familiar words to young learners.

Articles and prepositions are high-frequency function words that are found in the 1<sup>st</sup> frequency bands, which is not a surprise. Considering the total number of words, the English language does not include many high-frequency words – the limit is set at the 2<sup>nd</sup> frequency band – but these words make up a large portion of texts (Nation, 2001). The rest of the determiners and all the pronouns, adverbs, conjunctions, and most of the interjections of this material are from the 1<sup>st</sup> frequency band. “Welcome” is from the 2<sup>nd</sup> frequency band, but well in the scope of useful words for young learners to learn. The interjections “oh”, “shh”, and “wow” are not found from the first 20 frequency bands. *High five! 1–2* had a ‘food’-related task with “yuck” (10<sup>th</sup> frequency band) and “yum” (9<sup>th</sup> frequency band), but the words were accompanied by a big thumbs up and a thumbs down. While they are low-frequency words at least they are short and somewhat close to their Finnish counterparts “yök” and “nam”. “Hooray” in *Jump in!* is from the 9<sup>th</sup> frequency band but again close to the Finnish “hurraa”.

Understanding how varied the cognitive processes of young learners are would be helpful when designing items related to language learning and teaching (Butler, 2019). Some young learners may make more comparisons to their first language and sometimes the frequency band words are found in does not matter. “Vanilla” is from the 7<sup>th</sup> frequency band but as the word is almost identical to the Finnish equivalent “vanilja” that Finnish students will not have issues learning the word. Frequency bands range from the 1<sup>st</sup> to the 20<sup>th</sup>, which means that not all vocabulary presented in the material is suitable for young learners frequency wise. More than half of the words introduced are fortunately high-frequency words which means they can be found in the 1<sup>st</sup> and 2<sup>nd</sup> frequency bands and teaching them is justified.

### 5.3 Recycling of the Vocabulary in the Early Language Learning Material

The third research question is “How is the vocabulary recycled?”.

Not only is recycling of the vocabulary important when learning words, but practicing words in different ways supports gaining knowledge of the targeted foreign language (Yusupovna,

2023). Barclay and Schmitt emphasise that encountering and using words is extremely important if learners wish to learn those words (2019). Thus, studies provide support to the importance of recycling vocabulary. “The number of exposures necessary” to learn vocabulary varies; incidental and intentional learning plays a part in the equation along with the necessity of learning a word (Barclay & Schmitt, 2019). Optimal recycling introduces a word to young learners more than five times (Norberg & Nordlund, 2018, p. 467), but Nation suggests that after 12 times a word has been processed and retrieved enough times by the learners (2014).

If words are not recycled or revised enough, students are more likely to forget the words (Barclay & Schmitt, 2019). Recycled words appeared in three different ways in the material. The first type of recycling happened only within one chapter meaning that the word was introduced in a chapter where it appeared more than once, but it was not encountered anywhere else in the workbook. Introducing a word only once happened for example through songs, listening exercises, and pictures. The second type of recycling happened when the word appeared in a chapter and it appeared in it either once or more than once, but it also appeared in the extra page relating to the chapter at the end of the book. This means that the words were only met in the certain type of context. The third type of recycling was when words appeared in several main chapters, and at times also in the extra pages at the end of the workbooks. Vocabulary was also recycled in the self-evaluation tasks at the end of the main chapters in *Come with me! 1* and *2*, and *Go!*. Self-evaluation tasks are not uncommon in learning material (Butler, 2019, p. 486) and offer an opportunity to have young learners recall vocabulary that was taught in the chapter.

“Frequency of exposure has been found to correlate with acquisition for both incidental and intentional learning” (Barclay & Schmitt, 2019, p. 807). While words in the workbooks are mainly encountered intentionally, some exercises in the workbooks offer a chance of slightly incidental exposure as students are given different choices to make. Luckily studies prove that intentional learning does not involve as many numbers of exposure as incidental learning does (Barclay & Schmitt, 2019). While learning itself in a classroom setting is intentional, so should the recycling of vocabulary in the learning material be.

Hirsh and Nation have understood that people do not use English words equally (1992) and the corpus formed for this study proves that frequency wise words are not used in equal ways. All the words in the material are not recycled, as some of the vocabulary is presented only once in workbooks. Of the words that were categorised as nouns, 96 words appear only once in the whole material meaning that while only one workbook introduces the noun, it also appears only once in that workbook. 27,20% of nouns appear in only one of the workbooks meaning that they are not recycled at all, as there were 353 different word families of nouns.

The individual most recycled noun was “school” in *Go!* where it was encountered 34 times. While it was encountered plenty enough, it appeared in chapters 7, 8, 10, 11, and 12, which all are almost consecutive chapters. The number of times a word is encountered is not the only factor that defines whether words are learned, but the pattern of revision impacts learning. Spaced repetition offers a chance for better learning, revising shortly after the word is first learned is important after which it is not as important, but recall is still required (Barclay & Schmitt, 2019). If each word family is counted separately in each workbook, there are 607 word families of nouns in the material. Out of these 607 word families of nouns, 301 word families appear more than four times which is what Norberg and Nordlund describe to optimal recycling for young learners (2018).

Personal pronouns were the most recycled pronouns in the material, the “I” word family was recycled the most in each of the five workbooks, followed by “you” and the “we” word family. Learning personal pronouns is important as young learners need words to talk about themselves and the people around them. “He” and “she” were not recycled as much, which could be because it is easier to understand the concept of “you” and because Finnish does not use gendered pronouns. Demonstrative, interrogative, and relative pronouns are recycled poorly in each workbook, with “what” being a positive exception excluding *Come with me! 2* where it appears only twice. The pronouns, excluding personal pronouns and “what”, do not appear in core vocabulary and only in a few tasks which makes recycling them more difficult through the applicable tasks.

17 verbs were not recycled at all, which means that those words were only encountered once. As forgetting a word is most likely after it has been introduced, repetition and a chance for

the learner to retrieve the word is necessary (Barclay & Schmitt, 2019). If those 17 verbs are excluded from the data and each word family in each workbook is counted, 50 verbs were recycled less than five times. Verbs such as “clap” and “stomp” in *Jump in!* were only encountered in the song ‘If You’re Happy and You Know It’ and while both verbs appear only three times, the repetition may be enough for young learners to learn the verbs as they are accompanied by the movements of clapping and stomping. Verbs that are core vocabulary in the workbooks are recycled well as are the declinations of the verb “be”.

Each workbook introduced several adjectives that were only presented once. *Come with me! 1* did not have any adjectives other than those that were related to the names of the characters and those adjectives (“brave”, “curious”, “active”, “kind”, “calm”) were introduced only once in the front cover of the book. The other adjectives that appeared only once and thus had no recycling are adjectives mainly found in the covers of the books, outside of the main chapters, and in pictures. The only word that was spelled in two different ways within a workbook was *Jump in!* using the written form of the adjective “gray” (American English) once instead of “grey” (British English) that the workbook had used in every other case. As the concept of word families may be difficult for students to understand, publishers should stick to one type of spelling per word type to avoid confusion in young learners.

Colours were the most recycled adjectives in four books but *Come with me! 2* recycled each colour less than five times, “black” and “red” appeared only once. The reason for their poor recycling is most likely due to the fact they are recycled two to 12 times in the series’ first book. The most recycled adjectives in *Come with me! 2* are the opposites “happy” and “sad” instead.

Eight of the material’s 21 adverb word families were recycled. Out these eight adverbs, “not” appeared the most in the material and was also recycled the most throughout the different books, appearing in seven different chapters in both *Go!* and *Jump in!*. Adverbs such as “after” in *Go!* and “really” in *Jump in!* are recycled but only in the chapter the words are first introduced in, which is not the most optimal way of spaced repetition. As both *Jump in!* and *Go!* are books used in the first 2 years of primary school the pace going through the material may be so slow that spaced repetition actualises well. 18 adverbs, some overlapping, in the

different workbooks appear only once and these words are not recycled at all. As the number of non-recycled adverbs is so high, using them is questionable.

The only highly recycled preposition is “to” in *Go!* as it appears 33 times in the book. The recycling of “to” in other workbooks is much smaller as it appears less than five times in each of the other books. “To” was recycled in *Go!* mainly in the phrase “Nice to meet you!”, in a song and a speaking exercise. The preposition was also recycled throughout the book in five different chapters, including the first and last chapters. The other prepositions were recycled much less throughout the workbooks and what appeared only once in one book (e.g., “for” in *Come with me! 2*) was recycled 8 times in another (“for” in *Go!*). The differences in the number of repetitions in the workbooks differs because of the different types of tasks, texts, and songs in the books.

As there are only a few conjunctions they are recycled well in the material. “But” was used only once in *Go!* so it could have been left out altogether. “If” appeared three times in *Go!* and “or” four times between two workbooks, but only once in *Come with me! 1*. “And” was the most recycled word in the material but its use differed between workbooks as *Jump in!* introduced it 55 times and *Go!* 53 times, but the *Come with me!* series seven times total.

Overall, the best recycled interjections were the opposites “yes” and “no”. *Go!* is the only workbook where “yes” and “no” are recycled the same number of times (23), in *Jump in!* “no” is recycled once more (14) than “yes” (13) and in the other three workbooks “yes” is recycled more (13–8 in *High five! 1–2*, 12–10 in *Come with me! 1*, and 17–16 in *Come with me! 2*. “Yes” and “no” are recycled often as there are various discussion tasks with examples on how to answer questions that include either “yes” or “no”. The word family “hello”, which includes “hello” and “hi”, is recycled well in each workbook to with the smallest number of recycled words is seven in *Come with me! 1*. “Bye” was recycled 22 times in *Go!*, but surprising only six times in *Jump in!* and once in the other three workbooks.

Articles are the most recycled words in the whole material as they appear in the most contexts, especially before core vocabulary (e.g., “a teddy bear”, “a shirt”, and “the guitar”). There are only a few very poorly recycled determiners at all – “their” (1), “six” and “ten” (2)



in *Come with me! 2*; “more” (1), “eight”, “nine”, and “ten” (2) in *Come with me! 1*, and “many” and “what” in *Jump in!*. Numbers are recycled mainly more than four times each, but “seven” in *Come with me! 1* and “four” and “nine” in *Come with me! 2* are recycled only three times. Some numbers are recycled more than three times as much as others (e.g., “one” (31 times) and “eight” (8 times) in *Jump in!*). Numbers appeared in tasks that included mathematics such as “eight - three” and students are expected to write the answer in English next to the equation. Possessive determiners are recycled more than five times in all cases except “your” in *Come with me! 1* where it appeared only twice.

In conclusion, some words are recycled very well in the material while others appear only once. When words are not recycled and appear only once in the material their necessity in the early language learning workbooks should be thoroughly considered. All core vocabulary in the five workbooks is recycled which is the most important. Some of the recycled items only appear in one chapter which is not the best case, but better than not recycling words at all.

## 6. Conclusion

The aim of this study was to find out how do *Come with me! 1*, *Come with me! 2*, *Go!*, *High five! 1–2*, and *Jump in!* promote foreign language learning through vocabulary. Early language learning is not effective only because of the head start to language learning but rather if it can achieve the necessities that teaching young learners requires (Nikolov & Djugunović, 2019). The research questions give sufficient insight on the vocabulary of the material to answer the aim of this thesis.

Nikolov and Djugunović state that “typically, young learners start with positive attitudes and are motivated” (2019, p. 584). This motivation young learners begin with should be kept throughout all foreign language learning not only because it is mentioned in the FNCC of 2014, but if students are motivated to work with the workbooks and the vocabulary introduced in them it validates the choice of vocabulary to some extent. The five workbooks revolve around similar topics and thus the words between the books do not include much variation. This does not give much room between young learners to have different topics of interest as these interests should be included in teaching foreign languages. Studies prove the significant importance of considering students’ differences (Nikolov & Djugunović, 2019). While this thesis focuses on teaching material, hopefully the students’ interests are more included in other aspects of teaching a foreign language than they are considered in the workbooks of this study.

Teaching language through vocabulary should result in success. If variables such as self-confidence or motivation (Nikolov & Djugunović, 2019) are not considered when designing teaching material, the students may be set up for failure and the material does not promote learning words in a sufficient way. Teaching low-frequency vocabulary to young learners, especially if the words are not recycled enough, cannot be generally justified. Another take on the topic is that learning low-frequency words is not unproblematic as learning a foreign language in a non-speaking country of the target language may have a very different culture. Culture is a part of the young learners’ lives and teaching material may include words that are high-frequency in the young learners’ first language but low-frequency in English or words that relate to the students’ culture but are low-frequency words in English.

Word knowledge is not only defined by understanding what a word means, but language learners should also be able to use it in a correct context (Barclay & Schmitt, 2019). Vocabulary taught in the material appears in the context it is chosen to be taught in, in this case along topics that young learners are expected to relate to and be interested in. Laufer and Ravenhorst-Kalovski mention *sight vocabulary* which encompasses understanding words “out of context” (2010, p. 16). As the core vocabulary of the five workbooks was presented in different topics, it cannot be concluded whether students would recall such words without a context or especially if introduced in another context. Learning vocabulary through different contexts and topics promotes language learning as it offers students varied words to learn.

The material presents numerous amounts of words for young learners to learn by providing core vocabulary in each chapter to learn along other words in the chapters which can be learned not fully incidentally but in a similar way. All parts of speech are included in the material, but not all domains are taught as core vocabulary as it focuses on nouns, verbs, adjectives, and numbers that are categorised as determiners.

Nation describes high-frequency words as the words found in the 1<sup>st</sup> and 2<sup>nd</sup> frequency band (1,000–2,000 words) and these words are learned especially through formal learning (2012). The material includes vocabulary from all frequency bands from the 1<sup>st</sup> to the 20<sup>th</sup>. Low-frequency words are not the first words that young learners should as they are only building their word knowledge of the English language. Thus, not all vocabulary in the material promotes foreign language learning as some words are truly too difficult and irrelevant for young learners to learn and from frequency bands that are very out of reach from the young students. The more words a part of speech had categorised under it, the more variation between frequency bands occurred and therefore nouns were found from the most different frequency bands.

Barclay and Schmitt discuss the unclarity of how word knowledge weakens in its various forms, but recognise the proof that not all decaying happens at the same speed (2019, p. 806). Recycling words aims to emit the issue of forgetting vocabulary, but unfortunately the material does not provide the best opportunities for students to encounter all words frequently. As some of the words that were recorded into the corpus appeared in pictures

and songs, there is a greater chance that the word was intended more of as an extra word rather than being in the main focus of theme, so such words were less likely to reappear later in the material. There is no correlation between the recycling of same words in different workbooks, which suggests that each word is only used when necessary and that recycling aspects probably have not been thought through. Some of the small numbers of recycled words in *Come with me! 1* can be justified when the recycling of the same word is better in *Come with me! 2*, but even spaced repetition needs a smaller retrieval span than recycling words during two different grades. Some of the vocabulary introduced is recycled efficiently which promotes language learning, but words that are presented only once have no sensible reason for appearing in the material proven by previous research on language learning and recycling.

As this thesis focused only on the vocabulary presented in the physical workbooks, further studies should include online teaching material which is often used to support teaching in the classroom. Online material also offers the possibility for young learners to practice the targeted foreign language at home and thus there are more possibilities for recycling words and active retrieval of vocabulary. Research should also focus on how well the vocabulary is learned, especially as the material introduces low-frequency words from the latter frequency bands of the BNC/COCA word family lists. As new teaching material appears, workbooks should be compared to what is already available and in use to define their relevance.

While this thesis focuses on the vocabulary presented in the five chosen workbooks, English is learned in various ways in class and learning should not only depend on the book that is used. Workbooks are an important part of learning a foreign language, but I have demonstrated that learning cannot solely depend on them in terms of vocabulary learning. Vocabulary in the workbooks is generally not recycled enough, which means young learners might encounter words they forget because they do not have the chance to retrieve it from their memory. That noted, how the words of the workbooks are used by other means of teaching is up to each teacher. As three of the books are designed for both 1<sup>st</sup> and 2<sup>nd</sup> grade it is also important to plan using the workbooks in a sensible way so that learning is most efficient for the young learners.

Using ready-made materials is understandable for teacher as teaching foreign languages for young learners is a new concept but teachers must carefully consider the contents of the teaching material they are using. Not everything in the workbooks is reasonable to teach to young learners as proved in the frequency and recycling sections of the discussion. The materials studied in this thesis are not unworkable, but each workbook has its strengths and weaknesses that should be considered when using them with young learners.

In conclusion, *Come with me! 1*, *Come with me! 2*, *Go!*, *High five! 1–2*, and *Jump in!* promote language learning by offering a number of different types of tasks, core vocabulary to learn, texts to read, songs to sing, games to play, and material to listen to. Different parts of speech are included and introduced in various settings, but the main groups introduced to young learners are nouns, determiners (because of articles), verbs, and adjectives. The material also introduces high-frequency words that are the most important for young learners to learn, but also several low-frequency words that do not promote language learning as well. Words are both recycled very well, but also very poorly or not at all. Words that are recycled properly promote language learning, unlike those words that are introduced only once with no proper meaning. The material provides 9,377 word tokens which amounts to as many opportunities for young learners to learn words, but the choice of words and how they are recycled should be more carefully considered in future material.

## 7. References

### Primary sources

Harjula, S., Heikkinen, M., Läykki, P., Pere, V., Manninen, H. & Turpeinen, E. (2021). *Come with me! 1*. SanomaPro.

Harjula, S., Heikkinen, M., Läykki, P., Pere, V., Manninen, H. & Turpeinen, E. (2021). *Come with me! 2*. SanomaPro.

Haukka, E. & Rantanen, M. (2016). *Jump in!* Otava.

Kalaja, M., Korpela, N., Kuja-Kyyny-Pajula, R., Mäkinen, J. & Pelli-Kouvo, P. (2020). *High five! 1–2*. Otava.

Kanervo, P. & Laukkarinen, A. (2016). *Go!*. SanomaPro.

### Secondary sources

Azim, M. U., Hussain, Z., Bhatti, A. M. & Iqbal, M. (2020). Recycling of Vocabulary in English Language Teaching: From Theory to Practice. *Epistemology*, 7(8), 88–102.

Barclay, S. & Schmitt, N. Current Perspectives on Vocabulary Teaching and Learning. In Gao, X. (Ed.), *Second Handbook of English Language Teaching* (pp. 799–819). Springer. [https://doi.org/10.1007/978-3-030-02899-2\\_42](https://doi.org/10.1007/978-3-030-02899-2_42)

Basic Education Act. 21.8.1998/628.

Bauer, L. & Nation, P. (1993). Word Families. *International Journal of Lexicography*, 6(4), 253–279. Oxford University Press.

Bland, J. (2015). Introduction. In Bland, J. (Ed.), *Teaching English to Young Learners: Critical Issues in Language Teaching with 3–12 Year Olds* (pp. 1–12). Bloomsbury.

Brysbaert, M., Stevens, M., Mander, P., & Keuleers, E. (2016). How Many Words Do We Know? Practical Estimates of Vocabulary Size Dependent on Word Definition, the Degree of Language Input and the Participant's Age. *Frontiers in Psychology*, 7(1116), 1–11. <https://doi.org/10.3389/fpsyg.2016.01116>

Butler, Y. G. (2019). Assessment of Young English Learners in Instructional Settings. In Gao, X. (Ed.), *Second Handbook of English Language Teaching* (pp. 477–496). Springer. [https://doi.org/10.1007/978-3-030-02899-2\\_24](https://doi.org/10.1007/978-3-030-02899-2_24)

Cunningham, U. (2011). *Growing Up with Two Languages: A Practical Guide for the Bilingual Family*. Routledge.

Decree of Basic Education Act. 28.6.2012/422.

Decree of Basic Education Act. 20.9.2018/793.

Demircioğlu, Ş. (2010). Teaching English Vocabulary to Young Learners via Drama. *Procedia Social and Behavioral Sciences*, 2(2010), 439–443.

Edukustannus. (2023). *Get Ready! 1–2*. <https://edukustannus.fi/alakoulu/get-ready/>

Finnish National Board of Education. (2016). *National Core Curriculum for Basic Education 2014*. Publications 2016:5. Next Print Oy.

Finnish National Board of Education. (2019). *Perusopetuksen opetussuunnitelman perusteiden 2014 muutokset ja täydennykset koskien A1-kielen opetusta vuosiluokilla 1–2. Määräykset ja ohjeet 2019:1a. PunaMusta Oy*. [https://www.oph.fi/sites/default/files/documents/perusopetuksen\\_vuosiluokkien\\_1-2\\_a1-kielen\\_opetussuunnitelman\\_perusteet.pdf](https://www.oph.fi/sites/default/files/documents/perusopetuksen_vuosiluokkien_1-2_a1-kielen_opetussuunnitelman_perusteet.pdf)

- Gardner, D. (2008). Vocabulary Recycling in Children's Authentic Reading Materials: A Corpus-based Investigation of Narrow Reading. *Reading in a Foreign Language*, 20(1), 92–122.
- Green, C. (2019). Enriching the Academic Wordlist and Secondary Vocabulary Lists with Lexicogrammar: Toward a Pattern of Academic Vocabulary. *System* 8, 2019(102158).
- Hirsh, D. & Nation, P. (1992). What Vocabulary Size is Needed to Read Unsimplified Texts for Pleasure? *Reading in a Foreign Language*, 8(2), 689–696.
- Hudson, R. (1994). About 37% of Word-Tokens are Nouns. *Language*, 70(2), 331–339.
- Hulstijn, J. H. (2012). Incidental and Intentional Learning. In C. J. Doughty, & M. H. Long (Eds.), *The Handbook of Second Language Acquisition*, (pp. 349–381). Blackwell Publishing.
- Laufer, B. & Ravenhorst-Kalovski, G. C. (2010). Lexical Threshold Revisited: Lexical Text Coverage, Learners' Vocabulary Size and Reading Comprehension. *Reading in a Foreign Language*, 22(1), 15–30.
- Loewes, S. & Reinders, H. (2011). *Key Concepts in Second Language Acquisition*. Palgrave Macmillan.
- Ministry of Education and Culture. (16.4.2018). Opetushallitus ja opetus- ja kulttuuriministeriö: Kieltenopetus alkaa jatkossa jo ensimmäiseltä luokalta – 4,3 miljoonaa valtionavustuksia jaettu. *Valtioneuvosto*. <https://valtioneuvosto.fi/-/1410845/kieltenopetus-alkaa-jatkossa-jo-ensimmaiselta-luokalta-4-3-miljoonaa-valtionavustuksia-jaettu>
- Nation, I. S. P. (2001). The Goals of Vocabulary Learning. In *Learning Vocabulary in Another Language* (pp. 6–22). Cambridge University Press.



- Nation, I. S. P. (2006). How Large a Vocabulary Is Needed for Reading and Listening? *The Canadian Modern Language Review/La Revue Canadienne des langues vivantes*, 63(1), 59–82.
- Nation, P. (1993). Measuring Readiness for Simplified Material: A Test of the First 1,000 Words of English. In Tickoo, M. L. (Ed.), *Simplification: Theory and Application* (pp. 193–199). SEAMEO Regional Language Centre.
- Nation, P. (2004). A study of the most frequent word families in the British National Corpus. In Bogaards, P. & Laufer, B. (Eds.), *Vocabulary in a Second Language: Selection, Acquisition, and Testing* (pp. 3–13). John Benjamins Publishing.
- Nation, P. (2014). How Much Input Do You Need to Learn the Most Frequent 9,000 Words? *Reading in a Foreign Language*, 26(2), 1–16.
- Nation, P. (2021). Critical Commentary: Thoughts on Word Families. *Studies in Second Language Acquisition*, 43(2021), 969–972.
- Nation, P. & Anthony, L. (2017). Measuring Vocabulary Size. In Hinkel, E. (Ed.), *Handbook of Research in Second Language Teaching and Learning: Volume III* (pp. 355–368). Routledge.
- Nation, P. & Waring, R. (1997). Vocabulary Size, Text Coverage and Word Lists. *Vocabulary: Description, Acquisition and Pedagogy*, 14, 6–19.
- Nikolov, M. & Djigunović, J. M. (2019). Teaching Young Language Learners. In Gao, X. (Ed.), *Second Handbook of English Language Teaching* (pp. 577–599). Springer. [https://doi.org/10.1007/978-3-030-02899-2\\_31](https://doi.org/10.1007/978-3-030-02899-2_31)
- Norberg, C. & Nordlund, M. (2018). A Corpus-based Study of Lexis in L2 English Textbooks. *Journal of Language Teaching and Research*, 9(3), 463–473. <http://dx.doi.org/10.17507/jltr.0903.03>

- Nordlund, M. (2016). EFL Textbooks for Young Learners: A Comparative Analysis of Vocabulary. *Education Inquiry*, 7(1), 47–68. <https://doi.org/10.3402/edui.v7.27764>
- Nordlund, M. & Norberg, C. (2020). Vocabulary in EFL Teaching Materials for Young Learners. *International Journal of Language Studies*, 14(1), 89–116.
- Otava Oppimisen palvelut. (2023). *Skylight*.  
<https://oppimisenpalvelut.otava.fi/oppimateriaalit/luokat-1-6/skylight/>
- Pietilä, P. & Lintunen, P. (2014) Kielen oppiminen ja opettaminen. In Pietilä, P. & Lintunen, P. (Eds.), *Kuinka kieltä opitaan* (pp. 11–25). Gaudeamus Oy.
- Rixon, S. (2019). Developing Language Curricula for Young Language Learners. In Gao, X. (Ed.), *Second Handbook of English Language Teaching* (pp. 277–295). Springer.  
[https://doi.org/10.1007/978-3-030-02899-2\\_15](https://doi.org/10.1007/978-3-030-02899-2_15)
- Skinnari, K. & Sjöberg, S. (2018). *Varhaista kielenopetusta kaikille. Selvitys varhaisen ja vapaaehtoisen kielenopetuksen tilasta sekä toteuttamisen edellytyksistä kunnissa*. Soveltavan kielentutkimuksen keskus.
- Weninger, C. (2018). Textbook analysis. In Chapelle, C. A. (Ed.), *The Encyclopedia of Applied Linguistics*. Wiley & Sons. DOI: 10.1002/9781405198431.wbeal1489
- Yusupovna, K. D. Methodology of Teaching Foreign Languages. *International Journal of Education, Social Science & Humanities*, 11(1), 519–524.  
<https://doi.org/10.5281/zenodo.7542709>