



University
of Stavanger

The Faculty of Arts and Education

MASTER THESIS

Study program: Literacy Studies – Master’s
Degree Programme

Spring term, 2019

Open

Author: Martyna Onyszko

.....
(signature author)

Supervisor: Torill I. Hestetraet

Engelsk tittel: Learners’ English Vocabulary Size in Upper Secondary School A study of
receptive and productive vocabulary size in Vg1 in Norway.

Word of reference:

Vocabulary size, vocabulary measurement,
lexical richness

Pages:108.....
+ attachment/other:117.....

Stavanger,3/5 - 2019.....
date/year

Acknowledgements

First and foremost, my thanks go to my supervisors, who have made submitting this thesis possible. I would like to thank Torill I. Hestetraet, for the constructive feedback, inspiring discussions and guidance throughout the process, as well as willingness to share her knowledge with me. My thanks to Ion Drew for the very constructive language feedback.

The students who participated in the study deserve a thank you, too. Without them, the study would have not been possible. They are the reason why researching language is meaningful, so that they can be provided with proven methods and learning.

Finally, I would like to thank my loving family and friends, who provided me with strength to finish this project.

Martyna Onyszko

Sandnes

May 2019

Abstract

The purpose of the project is to examine the pupils' vocabulary size. 100 students from an upper secondary school in Norway participated voluntarily in the study. In order to measure the receptive vocabulary size, the students have taken an online vocabulary size test, which estimated how many word families they can. To examine the productive vocabulary size, students have submitted written work as part of English teaching. The texts have then been submitted to an online program, Vocabulary Profiler, Classic Edition. In this way, lexical richness, measured by lexical variation, lexical sophistication, lexical density, as well as collocations, lexical errors and use of academic vocabulary could be investigated.

The goal and motivation behind the project was to measure how many words students in an upper secondary school compare it with results for lower levels, and compare with the number of words one should be able to perform various activities in English (e.g. reading a book or viewing one TV series). In addition, investigating the students' use of academic vocabulary, as described in the Academic Word List, as well as the general vocabulary, cf. General Service List.

The results show that the average receptive vocabulary size of Vg1 students in an upper secondary school in Norway is 8,338 word families. Average productive vocabulary size is 4,769 word families. The students rely heavily on high-frequency words when writing in English. Mid- and low-frequency words are used to a limited extent. Collocations and academic vocabulary are rarely used.

Table of Contents

Acknowledgements	I
Abstract	II
List of abbreviations	VI
List of tables	VII
List of figures	IX
1. Introduction.....	1
1.1 Aim and scope	1
1.2 The LK06 National Curriculum and the competence aims	2
1.3 The Common European Framework of Reference for Languages (CEFR)	3
1.4 Research questions.....	5
1.5 Outline of the thesis	6
2. Theoretical background.....	8
2.1 Definitions of terms used in vocabulary research.....	8
2.2 The nature of a word.....	12
2.3 Historical trends in second language vocabulary instruction	13
2.4 Exposure to English as a lingua franca	15
2.5 Communicative Language Teaching (CLT).....	17
2.6 Receptive and productive vocabulary size	19
2.7 Vocabulary and reading.....	21
2.8 Vocabulary and writing	23
2.9 Assessing the quality of vocabulary knowledge.....	25
2.10 Teaching language.....	27
2.11 Teaching vocabulary.....	28
2.12 Incidental and deliberate language teaching.....	30
2.13 The Academic Word List.....	32
2.14 General Service List	33
2.15 Norwegian studies on vocabulary	34
2.16 International studies on vocabulary.....	36
3. Materials and methodology	41
3.1 Introduction.....	41
3.2 Informants and data material	42
3.3 English Vocabulary Size Test Classic Edition.....	44
3.4 Vocab Profiler (VP), Classic Edition.....	46
3.4.1 Lexical variation.....	48
3.4.2 Lexical sophistication.....	49

3.4.3 Lexical density	50
3.4.4 Lexical errors	50
3.4.5 Collocations	52
3.5 Ethical issues.....	52
3.6 Reliability and validity issues.....	53
3.7 Possible limitations of English Vocabulary Size Test Classic Edition	55
3.8 Possible limitations of Vocab Profiler, Classic Edition.....	56
4. Results	59
4.1 Introduction.....	59
4.2 Receptive vocabulary	60
4.2.1 Receptive vocabulary, Group A.....	60
4.2.2 Receptive vocabulary, Group B	61
4.2.3 Receptive vocabulary, Group C	62
4.2.4 Receptive vocabulary, Group D.....	63
4.2.5 The combined results of receptive vocabulary size	64
4.3 Productive vocabulary profiles, Group A	64
4.3.1 High scoring students' vocabulary profiles	65
4.3.2 Average scoring student's vocabulary profile	68
4.3.3 Low scoring student's vocabulary profile.....	69
4.4 Productive vocabulary profiles, Group B.....	71
4.4.1 High scoring student's vocabulary profile, Group B.....	71
4.4.2 Average scoring student's vocabulary profile, Group B.....	73
4.4.3 Low scoring student's vocabulary profile, Group B.....	75
4.5 Productive vocabulary profiles, Group C.....	76
4.5.1 High scoring student's vocabulary profile, Group C.....	76
4.5.2 Average scoring student's vocabulary profile, Group C.....	78
4.5.3 Low scoring student's vocabulary profile, Group C.....	79
4.6 Productive vocabulary profiles, Group D	80
4.6.1 High scoring student's vocabulary profile, Group D	81
4.6.2 Average scoring student's vocabulary profile, Group D.....	82
4.6.3 Low scoring student's vocabulary profile, Group D	83
4.7 Lexical richness.....	85
4.7.1 Lexical variation.....	85
4.7.2 Lexical sophistication.....	86
4.7.3 Lexical density	88
4.7.4 Lexical errors	88

4.7.5 Collocations	89
4.7.6 Academic Vocabulary (AWL)	91
4.8 The relationship between receptive and productive vocabulary knowledge	93
4.9 Brief summary of the results	94
5. Discussion of the results	96
5.1 Introduction	96
5.2 Receptive vocabulary size	96
5.3 Productive vocabulary size	98
5.4 Teaching implications and recommendations	101
5.5 Limitations of the study	103
5.6 Suggestions for further research	104
6. Conclusion	107
7. References	109
8. Appendix	113
8.1 Competence aims	113
8.2 Request for Participation in the Research Project	114

List of abbreviations

- AWL – Academic Word List
- BNC – British National Corpus
- CEFR – Common European Framework of Reference for Languages
- CLT – Communicative Language Teaching
- EFL – English as a Foreign Language
- ESOL – English for Speakers of Other Languages
- GSL – General Service List
- IB – International Baccalaureate
- L1 – first language
- L2 – second language
- LD – lexical density
- LS – lexical sophistication
- LV – lexical variation
- NSD – Norwegian Centre for Research Data
- TTR – type-token ratio
- VP – Vocab Profiler
- VST – Vocabulary Size Test

List of tables

Table 1: Knowledge areas of language proficiency (based on Bachman, in Kamaradivelu, 2005, p. 22)

Table 2: Examples of activities which make use of various strands of language teaching and quality of attention (based on Webb & Nation, 2017 and Nation & Yamamoto, 2012).

Table 3: Needed vocabulary size for 98% coverage of the text, based on Nation (2006).

Table 4: The receptive vocabulary size of learners of English from various countries, and the amount of instruction they received (based on Laufer, 2000).

Table 5: Classification of lexical errors (based on Read, 2000, p. 206).

Table 6: The receptive vocabulary size of participants in research Group A.

Table 7: The receptive vocabulary size of participants in research Group B.

Table 8: The receptive vocabulary size of participants in research Group C.

Table 9: The receptive vocabulary size of participants in research Group D.

Table 10: The productive vocabulary text profile of one of the highest scoring participants in research Group A.

Table 11: The productive vocabulary text profile of one of the highest scoring participants in research Group A.

Table 12: The productive vocabulary text profile of the average scoring participant in research Group A.

Table 13: The productive vocabulary text profile of the lowest scoring participant in research Group A.

Table 14: The productive vocabulary text profile of the highest scoring participant in research Group B.

Table 15: The productive vocabulary text profile of the average scoring participant in research Group B.

Table 16: The productive vocabulary text profile of the lowest scoring participant in research Group B.

Table 17: The productive vocabulary text profile of the highest scoring participant in research Group C.

Table 18: The productive vocabulary text profile of the average scoring participant in research Group C.

Table 19: The productive vocabulary text profile of the lowest scoring participant in research Group C.

Table 20: The productive vocabulary text profile of the highest scoring participant in research Group D.

Table 21: The productive vocabulary text profile of the average scoring participant in research Group D.

Table 22: The productive vocabulary text profile of the lowest scoring participant in research Group D.

Table 23: Type-token ratio of the high-, average- and low-scoring students in the study.

Table 24: Lexical sophistication of the high-, average- and low-scoring students.

Table 25: Lexical density of the high-, average- and low-scoring students in the study.

Table 26: Types of lexical errors in the corpus.

Table 27: Types of lexical collocations used in the study (based on Gitsaki 1999, p. 59).

Table 28: Types of lexical collocations as categorized in the Academic Collocation List.

Table 29: Types of lexical collocations used by the participants in Group A.

Table 30: Types of lexical collocations used by the participants in Group B.

Table 31: Types of lexical collocations used by the participants in Group C.

Table 32: Types of lexical collocations used by the participants in Group D.

Table 33: The academic vocabulary represented in the available corpora, as listed in AWL.

List of figures

Figure 1: The layout of the Vocabulary Size Test used in the study.

Figure 2: The combined results of all participants' receptive vocabulary sizes.

1. Introduction

1.1 Aim and scope

The proposed thesis is a study of the receptive and productive vocabulary size in English of Vg1 students (first grade of upper secondary school, or year 11) in Norway. The target group consisted of 100 learners from academic studies in a upper secondary school in Rogaland, Norway. The students attended the first grade, meaning that they were 16 years of age, with some exceptions where learners were 17. The research is based on data material gathered using an online Vocabulary Size Test (VST) available at my.vocabularysize.com devised at the Victoria University in Wellington, which measures receptive vocabulary size, as well as Vocab Profiler (VP), a data-driven software available online, which measures productive vocabulary size. The students' essays will be submitted to Vocab Profiler Classic Edition, which was developed by Tom Cobb, a Montreal-based linguist and researcher.

Studying the kind of vocabulary instruction which provides the best outcome in L2 teaching has a long tradition of research. Often, it includes studies concerning the frequency counts of the words (Nation & Newton, 1997, p. 238). Both teachers and learners of English as a second language may benefit greatly from applying research-based vocabulary instruction, since it has been shown to promote vocabulary development. This entails that the students are awarded with very good returns for their hard work (Nation & Newton, 1997, p. 238). English as a school subject in the Norwegian school system is given reinforced attention due to its importance for further academic and professional achievements. The status of English as a lingua franca, with its presence in the media, entertainment and academia, provides additional motivation for Norwegian learners to master lexis.

Vocabulary is a clear indicator of how well second language speakers may communicate. Elements of that proficiency are lexical knowledge, fluency and usage (Levitzky-Avid & Laufer, 2013, p. 127). The basic building blocks of a language that words constitute are necessary to make meaning in the process of communication. Indeed, after a prolonged period of focus on grammatical correctness, vocabulary has received focus after having been recognized as an important area of language acquisition (Read, 2000, p. 1). Moreover, it takes time and effort for L2 learners to reach a vocabulary size and knowledge level with natural and

rapid word recognition. Here, the focus will be on measuring the level of vocabulary, rather than attempting to understand the process of learning.

1.2 The LK06 National Curriculum and the competence aims

The current National Curriculum was implemented in 2006 and is divided into two main parts: The Core Curriculum and the competence aims for each subject. The Core Curriculum expresses the need for adapting to the changing society and the responsibility educational institutions bear to prepare learners for the challenges of that society (Utdanningsdirektoratet, 2011, p. 5). The Core Curriculum is also supposed to ensure equal access to education. The document is based on the regulations from the Education Act and aims to ensure that the acquired skills and knowledge help the individual to perceive, participate, experience, and excel. For this competence to be safeguarded, the following aspects of a learner must be developed: the spiritual, creative, working, liberally-educated, social, environmentally aware and, last but not least, integrated human being (Utdanningsdirektoratet, 2013b). Throughout the Core Curriculum, the underlying values for educators are highlighted, including the premise that education should foster equality and solidarity, develop appreciation for beauty, and provide familiarity with the technical heritage and the cultural knowledge.

In 2017, a new Core Curriculum was passed by the Norwegian Parliament. At the time of writing, it has not yet taken effect, as it will be introduced into the schools together with the new subject curricula, which are currently being revised. For all intents and purposes, it is a continuation of values and principles vindicated in the previous document. However, the core values that are to create a fundament for teaching are more explicitly defined in the first section. Among these values are human dignity, identity and cultural diversity, critical thinking and ethical awareness, creative eagerness, engagement and the need to explore, respect for nature and environmental awareness and last, but not least, democracy and complicity (Utdanningsdirektoratet, 2018). Although the new Core Curriculum has yet not been implemented, the schools across the country are preparing to introduce it into their teaching practice. Moreover, the core values specified in the document are to be implemented within every subject at every level throughout the school system. This is to ensure a common value base for students everywhere in Norway.

The second part of the National Curriculum includes every subject taught in state and private schools in Norway on every level and entails the purpose, main subject areas, framework for basic skills, and competence aims a student should achieve after the second, fourth, seventh, tenth grade of lower secondary school, as well as the first grade of upper secondary school. The role of English as an international means of communication is highlighted in the English subject curriculum (Utdanningsdirektoratet, 2018). The basic skills that are to be developed are speaking, writing, reading, numeracy, and digital skills. They are integrated into the competence aims and are also seen as contributing to the development of those aims. The competence aims are general statements which include understanding, evaluation, usage, interpretation, production, and discussion of various linguistic and cultural issues regarding the target language.

The competence aims are divided into four main areas: Language learning, Oral communication, Written communication, and Culture, society and literature. Specific competence aims that address vocabulary development are included in the Oral communication and the Written communication main areas and are formulated in the same way: “Understand and use a broad general vocabulary and vocabulary related to your own education program” (Utdanningsdirektoratet, 2013a). The students in the present study are enrolled in a general studies program which aims to prepare them for further academic achievements at a higher level. Thus, vocabulary related to one’s own education program will include academic lexis connected to the various fields taught in school.

The National Curriculum is currently under revision. The main objective is to create programs that are more relevant for the future, with clear priorities and strong connections between the school subjects. The suggestions for new course outlines are open for feedback at the time of writing. The hearing closes 18.06.2019 (April 2019 – Utdanningsdirektoratet, 2019). Three core elements are suggested: Communication, Language learning and Meeting with English-speaking texts. The emphasized focus on language has as its aim to give students larger area of choice in communicative situation (Utdanningsdirektoratet, 2019).

1.3 The Common European Framework of Reference for Languages (CEFR)

The Common European Framework of Reference for Languages (CEFR) was published by the Council of Europe in 2001, which was named the European Year of Languages. As a reference

tool, it provides a basis for syllabuses, assessment forms, textbook contents, and other aspects of assessment within second language teaching. Moreover, levels of proficiency are clearly defined in the Framework, which ensures comparability and transferability within the European context. The Framework in itself is a product of a desire to create an environment which supports exchange, encourages language teachers and learners to cooperate, as well as raising awareness regarding language acquisition, the attitudes, methods and needs that influence the learning environment (Council of Europe, 2001).

An important part of the Framework is an analysis and assessment the learners' communicative competence. The components of communicative proficiency have been specified in CEFR (Skulstad, 2018, p. 47). These are divided into general competences and communicative language competence, where the latter consists of linguistic competences, sociolinguistic competence and pragmatic competence (Skulstad, 2018, p. 47). Some of those areas overlap with the areas of language proficiency, described in table 1 in section 2.10. The situational awareness is essential, as use of the target language must be both understandable, precise and feasible in the daily life of the individual. The multilingual aspects of language teaching, in addition to the preservation of cultural diversity in Europe, is one of the main goals.

The CEFR has laid the groundwork for the assessment of students' overall and individual skills and categorization of the assessment criteria for Norwegian teachers for the last two decades. When it comes to assessment, the document defines the kind of activities and sets of criteria that may serve as a reference. Different kinds of assessments are listed, including: achievement assessment, proficiency assessment, continuous assessment, formative assessment, summative assessment, and direct assessment (Council of Europe, 2001). Assessment criteria are an important aspect of the planning phase of second language teaching, and there is emphasis on the positive achievements rather than deficiencies. It is a common practice in Norwegian schools, that the teachers consult the assessment criteria with the students before a test. Such an approach allows the students to gain a better understanding of what is required from them.

Moreover, regarding vocabulary proficiency, the CEFR provides teachers with a common understanding when it comes to this particular strand of language. The lexical ability, which is defined as the skill to use vocabulary correctly, comprises of lexical and grammatical aspects of language. The lexical aspect includes fixed expressions and single word forms. The grammatical aspect includes the following: articles, quantifiers, demonstratives, personal pronouns, question words and relatives, possessives, prepositions, auxiliary verbs,

conjunctions, and particles (Council of Europe, 2011, pp. 111-112). Moreover, the CEFR provides the educators with an assessment framework which allows one to investigate both the vocabulary control and range. In the vocabulary range, level B2 closely reflects aims similar to those expected in the National Curriculum: “Has a good range of vocabulary for matters connected to his/her field and most general topics. Can vary formulation to avoid frequent repetition, but lexical gaps can still cause hesitation and circumlocution” (Council of Europe, 2011, p. 112), whereas the competence aim from the English curriculum is as follows: understand and use a wide general vocabulary and an academic vocabulary related to his/her own education programme (both for oral and written communication – Utdanningsdirektoratet, 2013a).

Since the CEFR has contributed to an international focus on both the linguistic and pedagogical aspects of teaching, it has been a welcome development in the field (Golden & Kulbrandstad, 2018, p. 257). However, the framework has also been subject to criticism worldwide. The numeric system reflecting the proficiency level is supposed to standardize the assessment, making it possible for learners across borders to study or work abroad. Yet, such an approach invites an understanding that language learning happens in a strictly linear manner, where the various phases are clearly separate from one another. Moreover, the framework is intended to be applicable from kindergarten until higher education. However, the research shows that it is most useful with adult learners (Golden & Kulbrandstad, 2018, p. 258).

1.4 Research questions

The underlying objective of the thesis is to determine the English vocabulary size and knowledge of the targeted first year upper secondary school students in Norway. The estimate of their vocabulary size entails both productive and receptive vocabulary size. The estimated scores of the participants’ vocabulary sizes, alongside the qualitative analysis of the students’ vocabulary, were then used to draw conclusions about the overall state of vocabulary proficiency of the learners, as well as conclude with implications for teaching.

The research questions are as follows:

1. What is the receptive English vocabulary size of the targeted students in Vg1 upper secondary school, academic studies?

2. What is the productive English vocabulary size of the targeted students in Vg1 upper secondary school, academic studies?

In addition, two sub-questions are addressed regarding the productive vocabulary size, divided into qualitative and quantitative analysis:

3. What characterizes the productive vocabulary size described in terms of lexical richness of the learners' written samples?
4. To what extent do the students use academic vocabulary, as presented in the Academic Word List (AWL)?

In the qualitative part of the analysis, the kind of vocabulary known to students is described, with attention to different frequency levels. In order to investigate to what extent academic vocabulary is used by the 1st grade students, their essays are checked for the lexis from both AWL and the General Service List (GSL). More specifically, what vocabulary is known to which students is described. Perhaps there is a connection between the productive vocabulary size and other measurements. Moreover, other characteristics of language usage are defined, such as lexical richness, type-token ratio, and level of sophistication. Other qualities of the students' writing involve the use of collocations, lexical errors and use of academic vocabulary as presented in the AWL. All the mentioned qualities will hopefully shed light on what characterizes the vocabulary knowledge of the 1st grade students of English in a Norwegian upper secondary school.

1.5 Outline of the thesis

The general purpose of this thesis is to study the receptive and productive vocabulary size of the upper secondary school students in Vg1 in Norway. The thesis is divided into seven chapters. This present chapter consists of an introduction, description of the main scope and motivation for the thesis, in addition to the policy documents that form a basis for teaching English as a second language in Norway. The research questions are also listed. The second chapter surveys the theoretical background, including a description of words as lexical units, the historical trends in teaching vocabulary, as well as available Norwegian and international research within teaching and acquiring vocabulary. A distinction is made between receptive and productive vocabulary knowledge, and links between vocabulary and other language skills are presented. The limitation of existing research concerning knowledge about the students'

vocabulary size and the nature of the used lexicon at the level of interest forms the main motivation behind the research.

In Chapter 3, the materials and methods applied to examine the two types of vocabulary knowledge are outlined. The reasons behind choosing the two specific types of online software are presented. Using these facilitates answering the main research questions, in addition to the sub-questions presented in the current chapter. The preparations following the data collection, alongside the discussion about the ethical issues and the validity and reliability of both methods, are included. Possible limitations and issues with using the software are also added. The results and main findings are presented in Chapter 4.

Chapters 5 and 6 are intended as an arena for the discussion of the findings and conclusion, with possible implications of the main discoveries, as well as a comparison with existing research of similar aims and scope. Here, the distinction between the receptive and productive vocabulary size is made apparent. However, comparison, contextualization and analysis of these two kinds of knowledge are made. In the final chapter, Chapter 6, the main findings are recapped, conclusions are drawn from the main findings, and suggestions for future research are provided.

2. Theoretical background

2.1 Definitions of terms used in vocabulary research

The current study is to a large extent a consideration of what vocabulary knowledge is and how it may be measured and described. To achieve that aim, one should define the terms that constitute lexical proficiency, starting with establishing what a **word** is. Indeed, both defining and counting words may not be as straightforward as it seems. Researchers may deal with **lexical units**, which are items that function as a single meaning unit, regardless of the number of words it contains (Schmitt, 2000, p. 2). Yet researchers operate with more specific units in order to count one's lexical abilities. The first one is a **type**, a unit of counting in which one counts words in the written or spoken text only the first time they occur. Secondly, there are **tokens**, a unit of counting whereby one counts every word form in a spoken or written text (Nation 2013, p. 7). Counting tokens leads to eventually registering many repetitions. Then there is the **type-token ratio** (sometimes also referred to as lexical variation): the relative proportion of types and tokens in a spoken or written text. It is a widely used measure of the language development of both native speakers and language learners to define variation (Read 2000, p. 18).

Lemmas consist of a headword and of its inflected and reduced (*n't*) forms. For instance, the words *run*, *runs*, *running* and *ran* all belong to the same lemma: *run*. Normally, the items included in a lemma belong to the same speech part. **Homographs** are single word forms that have at least two meanings that are so different that they obviously belong to two separate word families (Read 2000, p. 18-20). Thus, when encountered in a text, the homographs should be counted separately, even though many computerized vocabulary size tests do not account for these.

Words as units of language may be further distinguished into **grammatical (function) words** and **content (lexical) words**. The former represents words such as articles, prepositions, pronouns, conjunctions, and auxiliaries. They are considered more as items belonging to the grammar rather than the vocabulary and carry little specific meaning on their own. They serve more to provide meaning together with another word, especially to modify meaning, provide links, etc. The latter describes words such as nouns, verbs, adjectives, and adverbs. These may convey meaning in isolation. When researching a learner's vocabulary knowledge, it is the content words that are essential (Read, 2000, p. 18).

As indicated, defining what a word is does not necessarily constitute a simple task. For the purposes of this study, word families will be taken into account when establishing the learners' receptive and productive vocabulary size. A **word family** consists of a headword (for instance, 'represent'), its inflected forms (*represented, represents*), and its closely related derived forms (*representative, representation* – Hirsh, 2010, p. 369). Counting the word families in the **corpus** (a large collection of texts, either spoken or written – Webb & Nation, 2017, p. 7) is useful when trying to determine the receptive vocabulary size. This is due to the fact that the vocabulary size test used in this study operates with word families according to the frequency levels with which they occur in the English language.

Consequently, a **frequency level** is a way of categorizing word families according to the frequency with which they occur, divided into 1,000 levels. Therefore, the first 1,000 level consists of the 1,000 most frequent words in a language and the next level consists of the next 1,000 words that are a slightly less frequent, although still often occurring in communication (Webb & Nation, 2017, p. 7). For a learner, the more frequent words are of higher importance, as lack of knowledge of them may hinder communication to a large degree. Word families may be simply divided into **high-, mid- and low-frequency words** (Schmitt & Schmitt 2014, p. 486). The last category represents the smallest value for the learner in a communication situation. Knowledge of word frequency should be of significant importance for language teachers, as they should target explicitly the high-frequency words (Schmitt & Schmitt, 2014, p. 486).

This categorization of lexical knowledge also invites categories such as **general, technical** and **academic vocabulary**. Such a division suggests which words should be targeted and learned, based on the value for the learner. **General vocabulary**, often represented in the General Service List (GSL), shows the most frequent words of general usage, as discussed in more detailed in section 2.14. **Technical vocabulary** implies words only used within specific fields, for instance medicine, astrophysics or botany, whereas **academic vocabulary** contains lexis that is most likely to be used by students, researchers, and lecturers (Webb & Nation, 2017, p. 7). The latter category is closely described in section 2.12.

The categorization of vocabulary has another, very important function: it helps determine the **learning burden** – ‘the amount of effort required to learn it’ - of various lexical units (Webb & Nation, 2017, p. 25). This should be a central consideration while choosing the kind of vocabulary for a second language teaching. A few factors impact the learning burden of a word. For instance, similarity to the mother tongue will increase chances of utilizing

existing vocabulary knowledge. The more the learner is advancing in language acquisition of English, the easier it should become to learn and make correct assumptions about the language (Webb & Nation 2017, p. 26). In this case, learners who have Norwegian as their mother tongue have a considerable advantage, as it belongs to the same language family as English. Moreover, such linguistic proximity should enable **inferencing**, namely a learning strategy where the reader tries to guess the meaning using contextual clues (Nation 2013, p. 354). Other aspects affecting learning burden are irregular spelling, difficult pronunciation, unfamiliar grammatical patterns or collocations (Webb & Nation, 2017, p. 26). All these aspects should be considered when introducing new words. For instance, one should avoid introducing many technical words at a time or one should try to teach words from one word family in one class, so as to ease the learning burden and reinforce repetition.

The main notion that will be discussed and focused on is **vocabulary size** (often referred to as breadth of vocabulary knowledge), which may be defined as the number of words a learner may produce or understand (Hirsh 2010, p. 372). As there are two kinds of vocabulary size, the current study contains two research methods to shed light on both of them. A distinction is made between the **receptive** and **productive vocabulary size**. The former is the kind of vocabulary knowledge that involves perceiving the form of a word while listening or reading and matching it to its meaning. It is sometimes referred to as ‘passive knowledge’ (for reading and listening – Webb & Nation, 2017, p. 33). The latter kind of vocabulary knowledge involves using the form of a word while trying to express meaning through speaking or writing, when also trying to produce the correct form. It is sometimes referred to as ‘active knowledge’ (for speaking and writing – Webb & Nation 2017, p. 33).

While receptive vocabulary size is easily quantifiable, productive vocabulary size is a more complex matter to investigate. It deserves a more qualitative approach, where other lexical aspects are considered. A more detailed distinction between the qualitative and quantitative approach is discussed in Chapter 3. The writing the students have submitted has been checked for both lexical richness, sophistication, density and their knowledge of academic vocabulary and collocations. **Lexical richness** is often used as an umbrella term that is used for the characteristics measured by analytical statistics that describe language use. These components include the type-token ratio (a term for lexical variation, explained earlier in this section), lexical sophistication, lexical density, and number of errors (Read 2000, p. 200). **Lexical sophistication** is an aspect of range of expression, which includes the use of technical terms and jargon as well as the uncommon words that allow learners to express themselves precisely

in the target language. **Lexical density** refers to the ‘percentage of content words’, as opposed to grammatical words (Read, 2000, p. 196), in a running text, whereas a **collocation** is a combination of two or more words that seem to form a ‘relatively fixed expression’ (Read 2000, p. 21 - all these notions will be further discussed in Chapter 4).

The first criterion to be considered when researching vocabulary is **validity**. A test is valid if it measures what it is supposed to measure (Schmitt 2000, p. 166). If one conducts a vocabulary size test, the participants’ ability to determine the meaning of the word should be the outcome of the test. However, it may be argued that if the participants are provided with a context, reading comprehension rather than lexical proficiency is tested. Unskilled readers will score lower, due to their inefficient reading, not because of their small vocabulary size, rendering the test invalid. Many issues can determine validity: the design of the test, the helping tools the participants may use or the sampling (choice of respondents from the population – Dornyei, 2011, p. 27). A general rule is that the bigger the sample size, the more it allows for more representative results (Schmitt, 2000, p. 166).

There are two types of validity: **internal** and **external**. The first kind, also referred to as **measurement validity**, determines the quality of the test, and how well the variables of the test explain the causal relationship of what is desired to be measured (Dornyei, 2011, p. 51). If any other variable that is not taken into account in the design of the study actually causes changes to the investigated variable, then the internal validity of the test has been compromised. The latter kind of validity, also referred to as **research validity**, describes research that produces results that may be generalized to another sample or study. If the results are only applicable to the one, particular research, it is deemed externally invalid (Dornyei, 2011, p. 52).

However, the sample size may compromise the second major criterion in language testing, which is **practicality** (Schmitt 2000, p. 167). The desire to devise a test that is easily checked does not require the student to sit over extended periods of time (compromising the results, as the student would answer quickly just to finish the test), should be balanced against the size of a sample which will provide the researcher with enough information (Dornyei, 2011, p. 27). Obviously, one cannot check all the words in a language to estimate the vocabulary size, but rather have a sample of words, for instance from all frequency levels. The test should have a format which allows for quick and unproblematic answers and which does not demotivate the participant, causing fatigue. Such effects may compromise **reliability**. A test is reliable if the results may be replicated over time. In general, the more questions the test contains, the more reliable the results are. If the results may be replicated any time, the test has **perfect reliability**

(Schmitt 2000, p. 166). The discussion regarding choice of methods and their construct is continued in Chapter 3.

2.2 The nature of a word

To assess the vocabulary level of a group of learners, it is important to understand what a word is. It is possible to count every written word form on a piece of paper, thus counting *tokens* (Read 2000, p. 18). However, if we are dealing with *types*, all the words written in a text should be counted, with disregard to repetitions (cases where a word is used multiple times, but only counted as one occurrence of a word). Then there are *lemmas*: a headword with all its inflected or reduced forms (all the word forms of a lemma carry exactly the same amount of meaning, only with different suffixes). For example, if the following words are mentioned: *treat, treating, treats*, then if one focuses on lemmas, these will be counted as one word, with *treat* as the lemma. Finally, *word families* should be mentioned. Various derivations and inflections of a word are included in the word family. Thus, both *treaty* and *untreated* belong to the same word family as the word *treat* (Nation, 2013, p.8).

Our perception of vocabulary knowledge as single words needs to be redefined and enriched with a broader view of lexical terms. The distinction between tokens, types, lemmas and word families influences both the number of the words in the corpus, as well as the learning burden. The idea of learning burden was defined already in the early thirties of the last century. It is the amount of effort it takes to learn a new word (Webb & Nation, 2017 p. 25). The number of words will naturally be largest when counting tokens and types and will be greatly reduced when considering lemmas and word families. Likewise, the learning burden is reduced when one can learn the word base, and then only add affixes and suffixes according to one's needs.

An important implication to the amount of learning burden is mentioned in Zhang (2016). Words in the human mind are stored in a way that reminds of a mental lexicon, rather than a dictionary (Zhang, 2016, p. 47). This means that the words are stored in a highly interconnected way, just as in any other network. In terms of organization, the words appear similar in form, such as spelling or pronunciation, as well as based on meaning associations (Zhang, 2016, p. 47). This provides pedagogical implications in language teaching, since two central processes should be emphasized: network building (learning new words) and mapping (creating links between the words). Learning a new word includes learning the word itself, with its form and

meaning, as well as where it belongs in the already existing mental lexicon. The form of the word is learnt implicitly, whereas meaning is learnt in an explicit way (Zhang, 2016, p. 48). The pedagogical implications of mental lexicon have been well explained in Zhang (2016).

The following list is proposed by Nation (1990, p.31) to describe the notion of ‘knowing the word’: knowing the meaning, the written and spoken form, the grammatical behaviour, the collocations, register, associations, and frequency of the given word. For the purposes of the present thesis, word families will serve as a measure unit of the receptive vocabulary size test (see my.vocabularysize.com). For the productive vocabulary size test, both types, tokens and word families will be used as measures to analyse and give a fuller picture of the students’ knowledge of words.

When it comes to the receptive vocabulary size of various language users, numbers vary greatly due to various definitions of the term ‘word’. Schmitt (2000, p.3) reports that an English native speaker with higher education will have a vocabulary size amounting to 20,000-word families. An average native speaker adds 1,000 word families annually from childhood until adulthood. Sætevik (2018) concludes from the findings of his study of 8th and 10th graders in a Norwegian lower secondary school that the vocabulary size of the 8th graders was 6,000 families and that 10th graders knew approximately 6,600 word families. One of the aspects to investigate was how the vocabulary size in upper secondary school compares to that in lower secondary school. Another factor was whether the vocabulary size used by the students relates to, covers or is perhaps entirely different from the lexis registered on the Academic Word List (AWL) and General Service List (GSL), as well as whether it has any influence on academic performance.

2.3 Historical trends in second language vocabulary instruction

Historically, vocabulary instruction has been an undervalued field in second language learning (Zimmermann, 1997, p. 5). In the early stages of second language teaching, the Grammar Translation Method was favored, which prepared students for the tasks of translating classical texts. This resulted in students acquiring outdated structures and terminology. Classes would normally include teaching grammar rules and glossary lists. The method dominated language teaching until the 1920s (Zimmermann, 1997, p. 7). It was challenged by new ideas within

teaching that emphasized exposure to listening and oral skills with meaning conveyed directly through the target language, without the need for translation.

Thus, the Direct Method emerged (Schmitt, 2000, p. 12). Attention to explicit grammar instruction was considerably toned down. The main idea was to imitate how language developed naturally, first through listening and speaking, and finally by reading and writing. The main criticism of the method lay in its trivial attitude and oversimplification towards the similarities between the first and second language. As such, it was never fully accepted in European or American institutions, even though its roots go back as far as to the nineteenth century (Zimmermann, 1997, p. 9).

In addition to the abovementioned methods, the Reading Method was also in function before World War II. Its intention was to improve the reading skills, since few people travelled internationally at the time it was introduced in the early twentieth century. The greatest opportunity to encounter a foreign language was through reading. Moreover, it was the first method, initiated by Michael West, which emphasized the need for improving vocabulary skills to enable reading. Word-frequency lists were suggested, alongside a research-based approach to selecting the vocabulary strand of second language teaching (Zimmermann, 1997, p. 10). In many ways, the Reading Method marked the introduction of the scientific treatment of the lexical aspect of second language teachings worldwide.

However, during World War II, the shortcomings of all these methods were revealed, when military personnel were less than fluent in foreign languages. A new method was necessary, which reinforced the ideas of behaviorism and attention to proper language habits, which included memorization, oral drilling, and attention to sentence patterns (Schmitt, 2000, p. 13). Thus, the Audiolingual Method was born. This method assumed that exposure to language and good language habits would enrich the learner's language skills. No additional vocabulary training was added. However, the behavioristic approach was attacked by Noam Chomsky already in the 1950s, defining the end of the audiolingual era in foreign language instruction.

The renowned linguist Noam Chomsky sparked a vital shift with his publication *Syntactic Structures* in 1957. In the paper, Chomsky argued that an individual retained an internalized mental grammar, which is projected in one's use of language. In opposition to this line of thinking, Hymes published an article in 1972 on the concept of communicative competence (Zimmermann, 1997, p. 12). Most importantly, both contributions presented an opposing view of the formerly approved ideas of language learning as merely habit development. The focus

changed towards using language for communication purposes. With the attention on fluency and not accuracy, communicative methods currently share the goal of making the target language familiar to the learner. Hopefully, the scientific evidence and growing body of research on how language is actually acquired will allow vocabulary to remain in the center of the language classroom.

2.4 Exposure to English as a lingua franca

The growing importance of English as the global means of communication is unlikely to change in the near future. The historical factors help explain that development: first, the British expansion since the 16th century resulted in the British Empire covering large areas everywhere on the planet. Later, the American influence has helped founding the position of the language. This led to more language users having English as a foreign language (EFL), in addition to those that have it as a first or second tongue. It is used in the international context, within business, tourism, education, entertainment, such as the movie and music industry, and much more (McGarrighan & Rugesæter, 2018, p. 162). Moreover, Norwegian native speakers, as a small language community, value good English proficiency in all aspects of life: on a personal level, in career-related settings, in education and business (Hellekjær, 2008, p. 1). Such heavy dependence on English as a tool for international communication is undoubtedly accounted for when considering the students' motivation and importance of knowing the language well and communicating with it fluently.

Norwegian children are introduced to the English language as a school subject early on in their educational path. English was introduced into the Norwegian school system as a compulsory subject in 1959. As mentioned in section 1.2, English is not considered a foreign language in the school curriculum, like Spanish, French and German, but has its own, much higher status than any other foreign taught language at school. It is considered a second language, rather than a foreign one (Brevik & Hellekjær, 2017, p. 3). As mentioned before, the English language curriculum is heavily based on the CEFR. Yet, studies report a rather high proficiency of Norwegian students in English, which may be a result of extensive extramural exposure to the language (Brevik & Hellekjær, 2017, p. 3).

Indeed, by the time students start learning English at school (at first grade of elementary school when six years of age), they have already been exposed to the language in multiple ways. One reason is the fact that English-language shows are not dubbed, but rather subtitled

(McGarrighan & Rugesæter, 2018, p. 163). Moreover, other activities, such as reading the news, using social media platforms, such as Facebook and Twitter, and listening to music with English lyrics, all contribute to supporting the classroom education. Brevik and Hellekjær (2017) mention video games as a very important source of language learning (p. 3). The massive exposure to English has specific influence on language proficiency. According to McGarrighan and Rugesæter (2018), it leads to mostly incidental learning, especially understanding. The pronunciation of Norwegian learners, however, suffers from the same deficiencies as that of many years ago (McGarrighan & Rugesæter, 2018, p. 163). Brevik and Hellekjær (2017), on the other hand, report increased reading proficiency among learners (p. 3). Again, they refer to receptive skills, rather than productive ones. When it comes to vocabulary, a decent extracurricular contact with English is also significant. A recent study at the University of Stavanger concludes that the students' vocabulary size was positively affected by the amount of time spent using English-speaking media (Vold 2018, p. 70). Consequently, improvement in vocabulary size influenced the students' language proficiency in general.

There is evidence of high proficiency in everyday English among Norwegian native speakers. The considerable exposure has influenced the position of English as a second language in Norway (Hellekjær, 2019, p. 71). Students may develop high proficiency in daily use of English. Language instruction should thus focus rather on more ambitious aims, such as developing academic vocabulary, reading for purpose, or reading strategies that would help dealing with unfamiliar words (Hellekjær, 2008, p. 1). A study conducted on senior upper secondary school students has shown that they do not meet the requirements to achieve the level demanded for admission to higher educational institutions in either Great Britain or Australia. This points towards weak instruction on academic English proficiency, neglected vocabulary development and neglected reading strategies as well as excessive dependence on textbooks as opposed to extensive reading (Hellekjær, 2008, p. 16).

The large exposure to English of the Norwegian teenagers, the language's high status among foreign languages, as well as the emphasis on the communicative aspect of language teaching should be reflected in the National Curriculum. Yet, traditionally, the English subject curricula in Norway tended to shy away from using the concept of communicative competence (Skulstad, 2018, p. 57). This is mostly due to the fact that as a technical term, it might have been difficult to understand for the parents and students. In the current curriculum (2019), the communicative language skills and language and cultural competence are separate main areas of language learning (Skulstad, 2018, p. 57). This may pose even a more confusing formulation,

suggesting that there is a difference between communicative competence and cultural knowledge. The next section deals in detail with communicative language teaching (CLT).

2.5 Communicative Language Teaching (CLT)

Communicative Language Teaching (CLT), which will be the focus of this section, is generally concerned with the language needs of the students, their situation, preexisting competencies and expected future use of the language. While bearing in mind the real-life application of language, the method targets both form and function, allowing the student to use the language also outside of the classroom's walls (Kumaravadivelu, 2005, p. 91). This approach has its source in the functional theory of language, with a focus on practical use in communication situations. It has a distinct relevance for the current study, since vocabulary, which is the main field of interest here, is an essential element contributing to communication and comprehension in foreign language learning (Orosz, 2009, p. 181). CLT may be defined as a language teaching approach where the main aim is to communicate successfully in a certain context, where the focus is on providing the learners with skills needed for use dependent of context (Skulstad, 2018, p. 63).

Communicative competence is the most important aspect in L2 teaching and learning (Skulstad, 2018, p. 43). Any approach seeks a specific design of materials, activities and techniques (Richards & Rodgers, 2015, p. 86). Such an approach helps safeguard the learner's autonomy, individual needs, as well as focus on cooperation and interaction between the learners. The method is thus strongly connected to the Core Curriculum, which applies in all Norwegian schools.

As a direct consequence of English being a lingua franca, the contemporary society strongly depends on proficiency in that language as a means of access to both education, technology, career opportunities, and other arenas of development. For instance, fluency in English will ensure a student access to the syllabus in the area of interest, a possibility to travel abroad on an exchange program, or to connect with international students and lecturers. Indeed, there is an ever-increasing demand for students and workers who communicate freely in English in all contexts and situations (Richards & Rodgers, 2015, p. 83). Such development induces a need for teaching the language that will target the communicative skills and competencies.

There are several aspects of proficient communication in target language, and these have been specified as whether something is formally possible, whether something is feasible to implement, whether something is appropriate and whether something is actually done (Hymes, 1972 in Skulstad, 2018, p. 45). All the above-mentioned aspects build on an idea that real communication always has a purpose and a function (Skulstad, 2018, p. 50). Examples of communicative situations involve ordering food at a restaurant, engaging in a conversation with someone or asking for directions to a specific place. Learners are bound to use their competence to complete these activities effectively.

Another reason why the communicative approach has developed, is the fact that the language teaching profession is in constant evolution. As a counter-reaction to the existing approaches throughout the 20th century, described in more detailed in section 2.3, the communicative approach has evolved as a part of paradigm shift in language teaching. The communicative ‘revolution’ was also a reflection on the educational changes in Europe in the 60s and 70s of the last century (Richards & Rodgers, 2015, p. 84). The emphasis on communication, getting the message across, is still very much an alive concept in Norwegian schools, rather than perfectly correct grammar and pronunciation (McGarrighan & Rugesæter, 2018, p. 164). The ultimate aim of language teaching is, undeniably, communication.

As with any method, there are advantages and drawbacks of using CLT as the main instructional approach. Countries with a long tradition of teaching EFL, such as India, after experiencing the shortcomings of the Grammar Translation method or the Direct Method, found themselves in need of an approach that prepares students for the practical situations and purposes of language learning (Sreehari, 2012, p. 87). As an approach, CLT seems to equip learners with language skills necessary for everyday situations, such as applying for a job or asking for directions. Learning in a CLT classroom happens through facilitating meaningful activities in the target language. Through everyday language expressions, practical information is handled in the target language and authentic material is used. Both reading, listening and speaking skills are challenged regularly (Sreehari, 2012, p. 91).

Within CLT, words are to be used in their cultural context, both to lessen the learning burden (as discussed in section 2.1) and to avoid relying on translation method (Zimmerman, 1997, p. 14). In the planning stage of a CLT class, relative usefulness of vocabulary presented in the course is taken into consideration, with frequency counts as basis for choice of the right words (Zimmerman, 1997, p. 14). However, use of frequency lists has been shown to contradict one of the main concepts of CLT, which is the fact that, as in L1, vocabulary in L2 will develop

naturally, due to communicative exposure to target language (Zimmerman, 1997, p. 15). Nevertheless, the CLT recognizes the importance of fluency over accuracy and, with time, gave vocabulary attention and an esteemed status among language's elements.

In addition, the face-to-face aspect of many of the communicative tasks may help learners in various ways. The speaker may gauge from the reaction of the listeners whether they understand the content and adjust it accordingly. By providing a meaningful context of the activity, the learners may infer the meaning of the words that may be unfamiliar or receive help in negotiating the meaning from other students. By using the newly-introduced items repeatedly in a productive activity, the retention may prove to be higher (Nation & Newton, 1997, p. 244). Generally speaking, there are comprehensive reasons for why CLT may prove to be a well-balanced and optimal approach to language teaching. This answers to the reasonable assumption that the more the learner engages in activities where the new word is being used, the more likely this word will be retained (Schmitt, 2008, p. 338).

There are disadvantages to the approach as well. In order for the approach to prove successful, the teacher must provide varied, authentic material. The material should both stimulate the subject knowledge, as well as language skills. For the teachers to provide meaningful, practical skills, they should learn the individual needs, wishes, and the potential future areas of use of the language. Access to materials, textbooks, computers, and printers should be provided for the learners to make use of language in practical situations (Sreehari, 2012, p. 92). Otherwise, CLT may not prove to be an efficient approach.

2.6 Receptive and productive vocabulary size

As mentioned above, there is a long tradition for researching vocabulary size and the research available is substantial (Nation & Newton, 1997, p. 238). The value of knowing a learner's vocabulary size is significant, as it may facilitate assessing the vocabulary needs and material the learner should be exposed to (Nguyen & Nation, 2011, p. 87). This is because it is believed that its influence on the effectiveness of second language proficiency is significant.

The term 'vocabulary size' refers to the number of words an individual knows. The goal is to use a sample of words from different frequencies and check how many families the learner is familiar with (Read, 2000, p. 31-32). The measurement provides a measurable, individual characteristic, not very much unlike a person's age or weight. The variable is measured in a test

where one relies solely on the test takers themselves to demonstrate their knowledge. Obviously, this is only one of many challenges concerning the reliability and validity of the tests. These will be discussed in section 3.6.

Receptive and productive knowledge of a word involves different types of knowledge (Ozturk, 2015, p. 107). Receptive use may involve guessing of unfamiliar words, finding out the synonyms of the known words, whereas productive use may involve using the word in a way that is grammatically or pragmatically correct (Ozturk, 2015, p. 107). The distinction between receptive and productive vocabulary size leads to a number of implications. The methodology used for testing the respective kinds of vocabulary knowledge must be different. Moreover, L2 vocabulary learning research shows that receptive learning may result in both receptive retention as well as some productive retention. Productive learning, on the other hand, results in considerable receptive retention, as well as the fact that receptive and productive learning combined lead to similar levels of productive knowledge compared to only receptive learning. Additionally, the findings suggest that productive learning is considerably more difficult than receptive learning (Mondria & Wiersma, 2004, p. 79).

According to Levitzky-Aviad and Laufer (2013) productive learning should be given priority. Perhaps learners need more supervision while learning the productive kind of vocabulary, although with a better outcome than only receptive learning. The fact that productive learning is more difficult than receptive learning may perhaps indicate that the findings in this study will show a larger receptive than productive vocabulary size of the observed students. Another study reports that the use of vocabulary (productive knowledge) in a foreign language develops slower than the knowledge of the vocabulary (receptive knowledge), which suggests these expectations (Levitzky-Aviad & Laufer, 2013, p. 127).

The remaining issue concerning vocabulary size is the distinction between receptive and productive vocabulary knowledge. Receptive knowledge describes merely the ability to recognize a word and connect it to its meaning. The tests that determine receptive vocabulary size measure different word knowledge than those that examine productive word knowledge (Hirsh, 2010, p. 374). Productive knowledge of a word, on the other hand, relates to the practical ability of knowing when and how to use a given lexical unit. According to researchers, productive knowledge is indeed more extensive and requires greater familiarity (such as the word's collocations, for example) with a certain term (Read, 2000, p. 26).

Counting the numbers of words in a language has given birth to the idea of thresholds for effective language use. Moreover, vocabulary researchers are interested in determining the number of words needed for various tasks, as well as for examining the vocabulary size for use in various fields (Hirsh, 2010, p. 226). For instance, the 95% threshold has been determined as appropriate for the reader to be familiar with 95% of the running text, and still be able to gain an understanding of the gist (Hirsh, 2010, p. 227). Below that number, the reader has less chance to determine the meaning of the unknown words from the clues included in the context.

As a result, to measure receptive vocabulary knowledge calls for a different test than productive vocabulary knowledge. For instance, receptive knowledge of a single word in a word family may automatically mean receptive knowledge of another word in the same family. However, one may not assume that productive knowledge of one word in a word family automatically leads to productive knowledge of another word in the same family (Hirsh, 2010, p. 374). Indeed, one's ability to recognize a word is not equal to the ability to use a word correctly. Thus, two different tests are applied in this study to measure both types of word knowledge. For the receptive knowledge, word families are used, whereas the productive knowledge is measured with word type units. More details about the applied tests are discussed in Chapter 3.

The productive vocabulary size needs to be addressed in a different way. It may be measured quantitatively, using number of word families known by the student. Alternatively, the productive vocabulary size may be described quantitatively, using lexical richness. Lexical richness is an important measurement, as it is believed to have a direct link to both communicative skills, word recognition, and fluency of reading, which, in turn, influence comprehension (Schmitt, 2000 p.51) in addition to mastering the language as a whole.

2.7 Vocabulary and reading

It is useful to think of vocabulary in relation to reading. There is plenty of evidence that reading facilitates vocabulary acquisition both for native speakers and learners of English as a foreign language (Read, 2000, p. 45-46). Vocabulary acquisition through reading may be further divided into two categories: incidental learning through reading, or intentional instruction involving contextualized, reading-related tasks (Paribakht & Wesche, 2003, p. 175-177). Many factors need to be fulfilled in order for vocabulary acquisition through reading to be successful,

and it appears that a systematic approach consisting of both incidental and intentional learning ensures best results.

First of all, it is reported that learners possessing only limited vocabulary may benefit from repeated exposure through reading (Schmitt, 2000, p. 150). Intermediate learners and advanced learners with vast vocabulary knowledge will only have a chance to encounter the unfamiliar vocabulary through extensive reading of specialized literature. Extensive reading ensures the repetition of the new lexis, which is necessary for retention. However, the regularity and amount of reading are not the only requirements to ensure successful lexical enrichment, regardless of the learner's level. The strategy that a learner will most likely use when encountering an unknown word is attempting to guess the meaning from the context. In addition to rich context offering plenty of clues, the factors that will ensure successful guessing of meaning are the learner's background knowledge of the topic, plenty of local clues instead of global clues, the presence of similar words that may cause erroneous guessing, the presence of cognates that may trigger rightful guessing, and the learner's overall guessing skills (Schmitt, 2000, p. 153; Read, 2000, p. 53). Additionally, the percentage of known words in the text will determine the overall difficulty level of the script, as already discussed above in section 2.4.

Should all the factors lead to the reader's rightful guess, it is still not certain whether the new lexis will be retained. If there was minimal effort put into guessing the word (due to, for instance, overly rich context), then the shallow processing would ensure little to no retainability. However, sometimes the context may be so obvious that the reader does not need to guess the word to understand its meaning. Schmitt (2000) points out that this inverse relationship in 'guessability and retainability' (p. 155) should not, by any means, diminish the role of inferencing in vocabulary acquisition.

Indeed, Paribakht and Wesche (2003) report interesting findings. They found that the learners who were only learning new vocabulary through incidental learning acquired some words, but they often stayed at the recognition level (i.e. receptive vocabulary knowledge). Those learners who were supplied with additional vocabulary instruction along with reading activities, both learned more new vocabulary and extended their knowledge beyond simply recognizing the words. The results are explained by the fact that the learners in the second group were led through a prolonged, deeper mental processing that helped them remember and understand the new words better. This implies that incidental learning supplied with meaningful instruction ensures the most effective vocabulary learning and retention (Paribakht & Wesche, 2003, p. 196).

2.8 Vocabulary and writing

Lexical richness is an umbrella term, that is determined by more specific measurements, such as lexical density, variation and sophistication, as well as number of errors (Read, 2000, p. 200). As a lexical term, it is one of the factors that distinguishes a fluent writer from an amateur. However, the link between vocabulary size and writing is often overshadowed by an emphasis on reading (Schmitt, 2000, p. 155). Lexical richness should be a priority area in enriching learners' writing skills. Research reports that it was the errors connected to vocabulary that were graded as the most serious by university professors (Schmitt 2000, p. 155, Nation, 2013, p. 177). Language teachers should be aware of the importance of vocabulary knowledge when teaching, and this strand of language is indeed more important to a well-formed text than grammatical correctness. Moreover, it is the lexical errors that impede comprehension more than any ill-formed sentence or grammatical structure. In addition, productive vocabulary use in writing has been proven to have a positive effect on general language level (Levitzky-Aviad & Laufer, 2013, p. 128). Also, learners themselves reported a positive influence of vocabulary learning on their writing (Levitzky-Aviad & Laufer, 2013, p. 128).

A feature that distinguishes less knowledgeable L2 writers from advanced ones, or even native speakers, is the simplicity and lack of precision in their choice of words. A fluent user of the target language is bound to pick a less frequent, perhaps a technical or academic term, whereas the not-so-fluent learner will opt for a high-frequency word, imprecise, and a more general word (Schmitt, 2000, p. 155). It may not be because the learner does not know a better, more precise synonym. On the contrary, the word may be easily recognizable. However, the writer may not have gained the productive written control of that word. It cannot be emphasised strongly enough how important vocabulary size is for the aspiring L2 communicator. Making the correct word choices is what distinguishes a learner who has adopted the conventions of the target language (Nation, 2013, p. 178).

In this context, using a wide range of vocabulary should be prioritised, and the goal of expanding the students' vocabulary size, both the receptive and the productive knowledge, should be addressed intentionally and not left to incidental learning. One solution is encouraging the use of dictionaries, which allows checking the known words, as well as checking the more precise terms. Schmitt (2000) reports the neglect of dictionary knowledge and skills as a viable impediment in such an approach (p. 156). Thus, practice at dictionary use should also be given attention. This shows that in vocabulary acquisition, many skills are

acquired and nurtured through a constant and systematic approach and knowledgeable instruction on the teacher's part.

Yet, simply being familiar with a specific word is not sufficient for proper productive usage of language. The processing of 'chunks' or 'phrases' in the L2 is just as important for precise and effective expression. Gitsaki (1999, p.43) emphasizes the need for learning lexical phrases in the target language (phonologically coherent, repeatedly used expressions of at least two morphemes), also known as collocations. A collocation is described as a sequence of words that commonly come together in combinations of what seems to be fixed expressions (Read, 2000, p. 21). Examples of such lexical items are 'thick eyebrows', 'to collect stamps', 'to accept a proposal' (Gitsaki, 1999, p. 1). The term was introduced in 1957 by Firth (Gitsaki, 1999, p. 2). Research may suggest that both the theoretical and applied treatment of collocations as part of second language acquisition has been insufficient (Gitsaki 1999, p. 26). However, the importance of instruction on collocations involves more than just defining a word. The change in approaches to learning English has accentuated the need for attention to words not only in isolation. The grammar-translation method, for instance, may produce words that are semantically compatible, yet do not produce acceptable collocations. A student may be aware of the fact that a tea may be strong, producing a collocation 'powerful tea', which is by no means conventionally correct (Gitsaki, 1999, p. 27). The correct use of collocations is by no means reserved for native-like speakers. However, it does elevate learner's production in the target language to be able to use the collocations fluently.

Consequently, developing conventionalized language forms is significant for the learner's communicative competence. Collocational knowledge is also listed as one of the features that indicates a proficient language user (Levitzky-Aviad & Laufer, 2013, p. 129). Additionally, knowledge of collocations has been found to be helpful in vocabulary development and self-learning strategies, such as inferencing (Gitsaki, 1999, p. 29). The evidence that collocational errors constitute a large percentage of all mistakes committed by learners of English as L2, provides strong arguments that collocational acquisition is a difficult element in language learning. Even proficient English speakers make collocational mistakes and would often tend to use free lexical combinations rather than collocations (Levitzky-Aviad & Laufer, 2013, p. 130). Thus, proficiency in using collocations should be a representation of one's linguistic progress (Gitsaki, 1999, p. 29).

2.9 Assessing the quality of vocabulary knowledge

There are numerous reasons for why assessing vocabulary knowledge is an important element of a second language teaching. Vocabulary assessment may be used when placing the learners in groups of different proficiency levels (diagnostic tests). It is an appropriate part of language mastering, as it is crucial in reading comprehension (Schmitt, 2000, p. 163). Another reason concerns the attitudes of the students. Research suggests that for the learners, the contents are more important if they will be tested. Should vocabulary be an element that is never assessed, the students may feel tempted to ignore that part of the instruction (Schmitt, 2000, p. 163). Regardless of the reason, assessing vocabulary needs to be addressed, reflected upon, and analyzed.

Vocabulary acquisition may be discussed along many axes: lexical knowledge versus lexical use, passive and active vocabulary, and breadth and depth of vocabulary knowledge (Levitzky-Aviad & Laufer, 2013, p. 128). Vocabulary knowledge may be described quantitatively (breadth of knowledge), or qualitatively (depth of knowledge). Lexical knowledge is the type of data a learner has acquired, whereas lexical use is a representation of information, for instance during producing a text. For learners of English as an L2, lexical knowledge is usually more advanced than lexical use, as the information may not be easily accessible at all times (Levitzky-Aviad & Laufer, 2013, p. 128). Moreover, the receptive-productive distinction suggests that there is a difference between vocabulary knowledge where learners may retrieve a word in a communication situation (active knowledge), and those who can only do that when a given word is presented to them (passive knowledge – Levitzky-Aviad & Laufer, 2013, p. 128).

Until recently, vocabulary assessment has been focused solely on retrieving the meaning of the word (Schmitt, 1995, p. 88). Generally, the kind of lexical information desired will determine the type of test one will devise. There are three dimensions of vocabulary knowledge one may want to consider: the degree of mastery (partial/precise), the degree of various aspects of word knowledge, or a focus on receptive or productive knowledge (Schmitt 2000, p. 167). If the wish is to estimate the receptive size of a learner's vocabulary, a simple yes/no test, or a multiple-choice test, may be devised. The vocabulary may target either English lexicon in general, or a selection of vocabulary from the second language teaching. The number of words or word families obtained this way will estimate the breadth of knowledge (Schmitt, 2000, p. 164). For many purposes, this type of assessment may be enough (Schmitt, 1995, p. 93).

Indeed, such types of tests have been used since 1890. The main argument for such a long tradition must be the simplicity of the test, which many researchers have seen as a virtue. For a long time, it was believed that the decontextualized encounter with the targeted word allowed even the youngest learners, or the learners with low reading abilities, to show what they knew (Read, 2000, p. 88). The simple format of the test, which can be easily digitalized, is appealing to teachers who need to test many learners, and possibly expose them to many hundreds of items during a course (Read, 2000, p. 89).

A few dimensions need to be considered before devising a vocabulary size test: defining the units, selecting a sample, and defining the format of the test (Read, 2000, p. 89). All these factors should be addressed, as they may have direct influence on the results of the test. For instance, if the format of the test is multiple choice, the result may be dependent on the choice of distractors between the possible answers. They may either be closely related to the actual answer, or very far-fetched possibilities, making it easier for the participant (Read, 2000, p. 88).

A majority of classroom tests often belong to the abovementioned category where only the breadth of the vocabulary knowledge is tested, yet the growing awareness of its limitations forces language teachers to focus on testing that provides them with the possibility of measuring the quality of vocabulary knowledge: indeed, how well does the student know the words (word families)? Comprehensive lists of what word knowledge entails have been devised: the form of a word (spoken and written), the behavior of a word (grammatical and collocational), the frequency, stylistic constraints, and the conceptual meaning and associations the word has (Schmitt, 1997, p. 11). Tests that focus on this quality of vocabulary knowledge, measure depth of knowledge. A number of reasons confirm that measuring vocabulary depth is of value to language teachers. Being able to probe the participant's knowledge in more ways than simply the meaning of the word, measuring partial knowledge of the word, assessing various aspects of the word or assessing the quality of academic vocabulary of students are some of the rationales behind testing the depth of knowledge (Read, 2000, p. 92).

Regardless of the merits of measuring the depth of the vocabulary, its major limitation is that it can only give an apparent indication of the learner's word knowledge. In fact, it has been suggested that only the most common meaning of the target word is assessed, and the participants are only asked to determine a synonym of the target word (Read, 2000, p. 90). Additional or figurative meanings of the word are seldom tested in such settings. Indeed, the learner may have learned to associate the target word with one of its synonyms, thus scoring 'correctly' on a test, but without really knowing what either of the words mean (Read, 2000, p.

91). Obviously, such a test does not give the teacher reliable knowledge of the students' vocabulary skills.

2.10 Teaching language

Language, as a well-designed instrument, permeates every living aspect of our existence as humans (Kumaravadivelu, 2005, p. 3). The linguistic presence, as obvious as it is, in communicating, education, art, literature, legal systems and many more, is undeniable, yet seldom acknowledged. We consider language as a system of coherent elements, such as sounds, letters, syllables, words or sentences. For pedagogic purposes, a list of areas of language has been devised in order to group the various competencies:

Table 1: Knowledge areas of language proficiency (based on Bachman in Kumaravadivelu 2005, p. 22)

Organizational knowledge	<i>How utterances or sentences are organized</i>
Grammatical knowledge	<i>How individual utterances are organized</i>
	Vocabulary knowledge
	Syntax knowledge
Textual knowledge	Phonology knowledge
	<i>How utterances are organized to form texts</i>
	Knowledge of cohesion
Pragmatic knowledge	Knowledge of rhetorical or conversational organization
	<i>How utterances are related to the communicative goal</i>
	Knowledge of ideational functions
	Knowledge of manipulative functions
Sociolinguistic knowledge	Knowledge of heuristic functions
	Knowledge of imaginative functions
	<i>How utterances are related to language features</i>
	Knowledge of dialects
	Knowledge of registers
	Knowledge of natural or idiomatic expressions
	Knowledge of cultural references or features of speech

Indeed, the various areas of knowledge work together to form lexical ability. Similarly, Skulstad (2018) lists various subcompetences of communicative competence. Often, however, the lexical knowledge and the ability are treated as different aspects of language proficiency. No distinction should be made between knowing language and using it (Kumaravadivelu, 2005,

p. 23). This is why the current research will encompass both the receptive vocabulary size (simply knowing the word) and productive vocabulary size (knowing how to use it, for instance in a sentence or with a specific collocation). Moreover, the research design will draw from all of the knowledge areas listed in the table above, proving once more what a complex process language acquisition is. Grammatical knowledge, which entails knowing certain vocabulary, is just a small part of the learners' proficiency, whereas all the other knowledge areas need to be engaged when writing a text. In addition to grammatical knowledge, students must engage organizational, textual, pragmatic, and sociolinguistic knowledge.

Combining so many areas of knowledge into one language teaching can be challenging. Many factors and limitations must be accounted for by the teacher, such as the students' prior abilities, the potential future use of language, the learning environment, the optimal outcome of the teaching, and more. A principle that has proven to gain a wide range of followers, mostly due to its applicability, is Nation's four strands of language teaching (Nation & Yamamoto, 2012, p. 167). By dividing a second language teaching into four equally important parts, a teacher may easily approach the following issues, such as the kind of activities that should be devised most often, how can vocabulary be taught or how does a well-balanced course look like.

The four strands of language are meaning focused input, meaning focused output, language focus learning, and fluency development. The first ingredient, meaning focused input, involves acquisition through reading and listening. This is, to a large extent, incidental learning. Similarly, meaning focused output embraces incidental learning, although through different activities: speaking and writing. More explicit kind of attention to the language's features, such as grammar, spelling or pronunciation, is offered by the third strand, namely language focus. Finally, fluency development targets producing meaningful language utterances in a fair tempo (Nation & Yamamoto, 2012, p. 167). Generally speaking, this principle is helpful in designing a well-balanced course, where one strand (language focus) emphasizes the lexical aspects explicitly, while the three others target the communicative skills.

2.11 Teaching vocabulary

There is specific and significant importance assigned to learning vocabulary. Words, as the building blocks of language, are key to conducting activities such as listening, reading, writing,

and speaking. Introducing vocabulary is regarded as a central part of children's books, and any stage of education where the meaning of new terms is presented, explained, used, and repeated. The quantifiable nature of vocabulary makes it possible to count the number of words, word families, or any other lexical measurement that suits the purpose. Moreover, the lexical aspect of language may be classified according to the frequency of occurrence (high-, mid- and low-frequency words), categorized into levels (first 1000 words that are occurring most often, for instance), or area of use and specialization (general use, academic use, technical use) (Webb & Nation, 2017, p. 7).

Distinguishing three broad frequency levels has proven to be revolutionary for vocabulary learning, first of all because it is an objective and easily measurable notion (Nation & Anthony, 2013, p. 7). The high-frequency words are defined as the first 1,000-third 1,000 word families that occur most often in English (Nation & Anthony, 2013, p. 7). The first most frequent 2,000 words are included in the General Service List (GSL – discussed in detail in section 2.13). The mid-frequency words are the wide purpose words counted from the fourth 1,000-ninth 1,000 word families. Low-frequency words are defined from tenth 1,000 words and beyond (Nation & Anthony, 2013, p. 8). Adult native speakers and native-speaking teenagers starting upper secondary school are expected to know some of the words from the middle-frequency level. Even though it is the low-frequency words that constitute the biggest group (from the tenth 1,000 up to 25th 1,000), it is the high-frequency words and proper nouns that make up for over 90% of the running words in English (Nation & Anthony, 2013, p. 8). This should suggest relative need for the various levels of vocabulary.

There is a relationship between frequency of occurrence of a word and how probable it is for a student to know that word. Since the high-frequency words are encountered more often, the students are more likely to know the high-frequency words rather than mid- and low-frequency ones (Nguyen & Nation, 2011, p. 94). There are, however, clear benefits to knowing mid-frequency vocabulary, such as capability to deal with English for authentic purposes, or better reading comprehension (Schmitt & Schmitt, 2014, p. 495-496). Interestingly enough, there is a lack of a systematic approach to teaching mid-frequency vocabulary, even though there seems to be an understanding that the high-frequency vocabulary and low-frequency vocabulary is often addressed explicitly through topic-based input (Schmitt & Schmitt, 2014, p. 498).

Language teachers should, to the extent this is possible, use methods that are proven to provide the best learning outcome. There are reports of a long tradition of investigating

vocabulary learning according to the frequency with which words occur in language (Webb & Nation, 2017, p. 9; Nation & Newton, 1997, p. 238). Learners at their early stages of learning tend to be familiar with the most frequent words in English – more specifically, the first 1,000-word frequency level (the thousand words that, according to research, occur most often in English – Webb & Nation, 2017, p. 9). Focusing on acquisition of the highest-frequency words is said to yield good results (Nation & Newman, 1997, p. 238).

One should be careful, however, about assuming that a learner benefits from knowing as many words as possible. The incremental nature of vocabulary learning suggests that there are many other aspects of lexical knowledge to be fathomed so that a word may be considered ‘learned’. This entails, for example, both the collocations of the word, its grammatical forms, the spoken and pronounced form, and idiomatic expressions. This kind of knowledge is often denoted as the depth of vocabulary (Schmitt, 2008, p. 333). This measurement is helpful when distinguishing the productive and receptive vocabulary size: the research reports that the receptive lexicon is always larger than the productive one. In addition, there is a distinction between the spoken and written vocabulary size, with the latter usually being larger (Schmitt, 2008, p. 335).

There are many aspects that need to be accounted for when teaching language in general, and vocabulary specifically. The learner’s involvement with the word is reported to have a large influence on lexical acquisition. This notion may be broken down into three elements: need, search, and evaluation. Need shows in the necessity of knowing the word to solve a particular task or answer a question. Search is the act of actively trying to figure the word’s meaning, for instance by looking it up in a dictionary, or asking the teacher. The last element, evaluation, denotes the act of analyzing whether the given word fits the purpose and is the correct one (Schmitt, 2008, p. 338). Seeing as this model helps retain lexical elements effectively, one should bear in mind to devise tasks that involve all these actions from the learners’ side.

2.12 Incidental and deliberate language teaching

It may seem as if the vocabulary we learn is mostly a product of carefully designed courses and meticulously conducted activities. However, research reports that the majority of lexical acquisition happens while the learner is engaged in a language task, such as reading or listening, and the new words occur as they go along (Schmitt, 2000, p. 121). Implicit learning is defined as learning vocabulary as a by-product of an activity (Webb & Nation, 2017, p. 48). The words

are not targeted for deliberate pondering and learning their meaning, yet they need to be understood and familiar for the learner to understand a question or the gist of a sentence. Moreover, the amount of incidental lexical acquisition depends on the amount of input, and it belongs to the meaning-focused input strand of language teaching discussed above (Nation & Yamamoto, 2012, p. 167).

A large body of research provides evidence for what kind of acquisition incidental learning may benefit. As mentioned, the more input the learners experience, the more chance that the vocabulary incidentally processed will be retained (Webb & Nation, 2017, p. 48). Moreover, there is evidence that incidental learning is most useful when learning collocations, which are mostly picked up during reading. Thus, extensive reading should be a considerable part of a linguistic course. Besides, speaking and writing may account for the productive side of collocation learning, thus developing fluency. Here, both the meaning-focused output strand and the fluency development strand are included to facilitate the incidental vocabulary learning (Nation & Yamamoto, 2012, p. 168).

An opposite approach to incidental vocabulary learning is intentional vocabulary learning. Deliberate teaching involves conscious focus on a specific linguistic element, such as a word or a phrase (Webb & Nation, 2017, p. 62). In general, the major difference between the two approaches is the quality of attention required in certain activities. The following table provides an overview of activities, together with the kind of attention required and the strand of language teaching it covers:

Table 2: Examples of activities which make use of various strands of language teaching and quality of attention (based on Webb & Nation, 2017 and Nation & Yamamoto, 2012).

Strand	Activity	Quality of attention
Meaning focused input	Extensive reading graded readers	Incidental
	Listening to a recorded story	Incidental
	Browsing the internet in the target language	Incidental
Meaning focused output	Using a new word in an e-mail to a fellow student	Incidental
	Doing free writing using previously encountered words	Incidental
Language focused learning	Consulting a dictionary	Deliberate
	Studying using a workbook	Deliberate
	Memorizing words using flash cards	Deliberate
	Being taught words	Deliberate
Fluency development	Browsing websites in the target language	Incidental
	Watching familiar movies	Incidental
	Keeping a diary	Incidental

Learning a word in an activity has a much higher chance to be learned if it is focused on deliberately, rather than incidentally. Whereas incidental learning might be a time-consuming activity, and is largely dependent on quantity of input, deliberate language activities may perhaps account for one or two collocations, meanings, or associations of a given word. Indeed, one of the biggest challenges of deliberate language teaching is the use of words without providing a context (decontextualization), which may result in limited areas of use (Webb & Nation, 2017, p. 62).

Optimal second language teaching should provide more opportunities for incidental than deliberate learning (Webb & Nation, 2017, p. 62). However, the deliberate attention to linguistic features is only accounted for in one of the language teaching strands, i.e. the language focused learning strand. Thus, deliberate learning should not make up more than one fourth of second language teaching as a whole, nor any other activity (such as learning collocations – Nation & Yamamoto, 2012, p. 169). When choosing activities, one should consider their effectiveness. In addition, opportunities for switching between incidental and deliberate learning should be taken into account. For instance, incidental learning from reading a text may be reinforced by deliberate focus on specific collocations and phrases from the input (Nation & Yamamoto, 2012, p. 173).

2.13 The Academic Word List

The Academic Word List (AWL) was devised by Averil Coxhead as a part of her master's thesis at Victoria University of Wellington (Coxhead, 2000). It includes 570 words listed alphabetically, with an index that suggests their frequency (1 is the highest frequency, 10 is the lowest). The primary goal of the list was to help language teachers to prepare students for higher education. The list does not contain the first 2000 most frequent words in English, nor does it include proper nouns, Latin names, or narrow range words (Coxhead, 2000). As the students in the present study are in their general courses at upper secondary level, it is a relevant question whether the English course they are enrolled in covers the words from the AWL. In addition, the extent to which the students already use the vocabulary included in the AWL may show direction for the design and content of vocabulary instruction in the classes for the remaining part of the course. The AWL in the Vg1 English course is not part of the study. However, the vocabulary from the AWL that the targeted students use, is a part of the research.

The current study is based on already available research regarding word frequency studies (Nation & Newton, 1997 p. 239). Word frequency studies constitute the background for this thesis, especially the high-, mid- and low-frequency distinctions. The GSL mentioned above covers around 2,000 word families and almost 80% of the words in a running text (Nation & Chung, 2003). The AWL, on the other hand, contains all the words significant for academic purposes. The low-frequency words include all the words not present on the GSL, AWL, or other technical lists, and not the high-frequency ones (Nation & Anthony, 2013 p.8). Moreover, there are certain thresholds which a speaker should reach in order to be able to participate in specific activities, such as reading a novel or watching a children's movie. These are described in more detail in section 2.15.

2.14 General Service List

It is clear that when dealing with vocabulary size, precise information about the kind and frequency of vocabulary used in testing is important. Specialized vocabulary is widely used and compiled in glossaries and specialized term lists, and the Academic Word List may serve as an example. When it comes to high-frequency words, general service lists have usually been used both by language teachers and learners. The motivation behind this is that the most frequently used words are so crucial that learning them is time well spent, as they will be used and encountered often in the target language. In the decades preceding the possibility of conducting computerized tests, Michael West carried out research where the result was a list of the 2,000 most common words in the English language (Read, 2000, p. 160).

The GSL contains approximately 2,300 of the most common words in English, yet not the most frequent ones. The advantage of the GSL, even though already somewhat outdated, is that it provides its users with the meaning of the words, such as the necessity or the learning ease. It provides a large lexical coverage, as it has been devised using a corpus of five million words (Webb & Nation, 2017, p. 10). Michael West is indeed the father of the notion of frequency as a tool in L2 learning (Schmitt, 2000, p. 17). It is he that revealed that to increase readability, low-frequency words should be replaced with more commonly used lexis. By using less specific vocabulary, the length of a text is extended, which substantially reduces the percentage of new words occurring in the text relative to the number of words. In this way, the reader does not have to cope with many new words in a short time span.

Another advantage of the GSL is the fact that it not only provides a list of high-frequency words, but also a list of meaning senses. Thus, a teacher that does not follow any contextualized activity may still be provided with a context. The major shortcoming of the GSL is the fact that it is based on corpora collected before 1936 (Schmitt, 2000, p. 84-85). Needless to say, the register has changed a good deal since then. However, it is reportedly still the most highly used and quoted high-frequency word list (Zimmermann, 1997, p. 9).

2.15 Norwegian studies on vocabulary

Available research within the Norwegian EFL classroom often focuses on the link between vocabulary knowledge and reading comprehension. Many age groups have been granted relatively limited attention and that is where researchers may still contribute to the expansion of that field. Nordby (2007) investigated the vocabulary development of 183 children in the second grade of elementary school in Oslo. This age (8-year-olds) is of special interest, as it precedes the development of functional reading skills. It was concluded in the study that there is a relationship between the child's vocabulary depth, breadth, grammar, inference, and working memory (Nordby, 2007, p. 4). Grammar and inference seem to have the biggest influence on vocabulary depth and breadth. The complexity of the relationships, as well as the fact that they influence each other, confirms that vocabulary is not an isolated strand of language, but exists in interaction with other language components.

When it comes to the vocabulary size of older student groups, Sætevik (2018) has recently carried out a cross-sectional study of productive and receptive vocabulary in the 8th and 10th grades. After testing 134 students' vocabulary knowledge, it was concluded that the receptive vocabulary size of 8th graders was on average 6,000 word families, while that of 10th graders was 6,600 word families. Their productive vocabulary, on the other hand, was estimated at approximately 5,000-6,000 word families for both levels. Sætevik concluded that students may acquire roughly 300 word families annually, averaging 3,15 words per class. However, the enormous variation should be pointed out: the lowest score in the 10th grade when it comes to receptive vocabulary knowledge was 1,800 word families compared to the highest score, which was 15,200 word families (Sætevik, 2018, p. 86). This is a reflection of the vast variation of knowledge and prerequisites within a single class.

Concerning the older age groups, Skoglund (2006) carried out a comparative analysis of American and Norwegian students' vocabulary usage at the upper secondary level. The comparison comprised of both qualitative and quantitative analyses. The findings included the following: the native speakers produced more types of words than the Norwegian speakers, although the difference was not as significant as expected. The productive vocabulary size of the Norwegian upper secondary school students was estimated to average 4,000 word families. (Skoglund, 2006, p. 75). It was concluded that the Norwegian learners in general needed to expand their vocabulary in order to improve their writing skills. The results contrast strongly with Sætevik's (2018) findings described earlier.

In another study, Opdal (2017) investigated the ability of Norwegian students to distinguish between general vocabulary and academic vocabulary in English. By comparing a regular, general credit class of third grade students at upper secondary level and a class enrolled at the International Baccalaureate (IB), the findings were, as expected, that the students in the IB class were more aware of the different registers of language. They distinguished properly between formal and informal vocabulary and knew which belonged in a specific communication situation. As a result, they used academic vocabulary more often, more precisely, and more fluently than the students from the regular class. It should be noted at this point that the IB class is a full immersion programme, where all subjects are taught in English. However, the difference found by Opdal is not significant (Opdal, 2017, p. 47). Further research should be conducted in this field.

Skjelde (2015) has studied academic vocabulary used in textbooks for the English subject at Vg1 level, upper secondary school in Norway, in order to analyse its influence on vocabulary acquisition. Twenty one factual texts, extracted from three different textbooks, were analysed, all of them related to the competence aim from the area of Society, culture and literature: *discuss and elaborate on texts by and about indigenous peoples in English-speaking countries* (Utdanningsdirektoratet, 2013a). The findings suggest that academic vocabulary is used only to a limited extent in the textbooks. Consequently, if the students are to acquire that kind of lexical proficiency, they need to be supplied with input that is both more authentic and they need to use more academic vocabulary. Additionally, academic vocabulary is not targeted, as it is not explained in glossaries. As a main implication, Skjelde suggests that other input is used in addition to the textbooks which targets academic vocabulary at the upper secondary school level in Norway (Skjelde, 2015, p. 99-102).

2.16 International studies on vocabulary

The learning environment of a person trying to master a second language is dependent on so many factors that researchers still have not managed to deliver a clear description of what approach is the best concerning vocabulary acquisition (Schmitt, 2008, p. 329). One may think that this is due to the lack of research in the field. This is, however, not the case, as substantial research is now accessible to L2 learners, teachers and researchers worldwide. The large body of evidence has undoubtedly had an impact on teaching vocabulary. Yet, factors such as materials, methods, language proficiency, and background of the learners would suggest that there is no consensus as to how to teach language in the most effective way (Webb & Nation, 2017, p. 1). This section provides an outline of the vocabulary acquisition research, specifically with an emphasis on measures concerning vocabulary size and word frequency.

One line of inquiry is the aspect of word frequency, which has been the very foundation for selecting the appropriate lexicon for second language teaching. Paul Nation has been an extremely valuable contributor in researching the vocabulary and regularity of its occurrence and has divided it into five groups: high-frequency words, mid-frequency, low-frequency words, academic words, and technical words (Nation, 2011, p. 9). Nation (2007) emphasizes the necessity of determining the right vocabulary to be taught, merely by conducting a cost-benefit analysis of what will facilitate the students' language skills the most. Thus, high-frequency vocabulary, consisting of the 3,000 most commonly used words in English, is crucial to be effectively retained by the students. These words need to be taught deliberately. Low-frequency words, on the other hand, are so uncommon that there is no need to spend classroom time to learn them. Rather, the teacher should teach strategies to help learners to acquire these lexical items on their own (Hirsh, 2010, p. 373).

The word frequency approach has been a largely accepted concept and motivation behind choosing the target vocabulary. It is nonetheless a problematic issue. Schmitt and Schmitt (2014) point out that should the frequency counts be of any value to the second language teachers, the distinction between high- and low-frequency words should be meticulously determined (p. 485). Firstly, one often discusses the frequency levels in terms of 1,000 word families. Yet, the academic and technical vocabulary lists do not follow the same pattern. Thus, these word lists do not 'fill the gap' between the two frequency levels. Rather, they add another dimension of vocabulary register. Moreover, highlighting the mid-frequency vocabulary as a

middle ground should be addressed pedagogically (Schmitt and Schmitt 2014, p. 501). This remains to be investigated.

Interestingly, Schmitt (2008) also reports that teachers themselves cannot always rely on their intuition when it comes to dividing vocabulary into high- or low-frequency categories. Thus, the role of a researcher as the provider of correct and reliable information is crucial in facilitating vocabulary acquisition. In these terms, providing, for instance, frequency lists, or lists of academic or other vocabulary lists (such as the AWL and GSL discussed in previous sections), is certainly a step towards more effective language instruction (p. 333).

Another line of inquiry is how much vocabulary a language user needs. One of the main components of second language instruction is highlighting realistic teaching goals. Research reports that it is in fact a coverage of 98-99% (meaning that one word in every 50 is unfamiliar). This figure is somewhat more uncertain for spoken discourse than written (Schmitt, 2008, p. 331). The number of familiar words is crucial for determining whether the learner will achieve such coverage. Briefly mentioned in section 2.7, the following table provides a more specific summary of the needed vocabulary size in order to achieve 98% coverage of the material, compiled from the data collected by Paul Nation:

Table 3: Needed vocabulary size for 98% coverage of the text, after Nation (2006).

Activity	Number of word families
Reading a novel (<i>The Turn of the Screw</i>)	7,000
Reading a newspaper	8,000-9,000
Reading a novel (<i>Lady Chatterley's Lover</i>)	8,000-9,000
Reading graded readers (Level 3 Bookworm, <i>The Picture of Dorian Gray</i>)	3,000
Watching a children's movie (<i>Shrek</i>)	7,000 (but only 4,000 for 96,7% coverage)
Following unscripted spoken English	6,000-7,000

Considering 98% as the optimal coverage for an unhindered understanding of input, approximately 8,000-9,000 word families should be familiar to the reader, or 6,000-7,000 word families should be familiar to the listener. Still, this deems one in every 50 words to be unknown (Nation, 2006, p. 79). A successful language teacher should make use of the available research by targeting the high-frequency words over the low-frequency ones, trying to establish the

learner's vocabulary size and potential registers of usage in order to target the right vocabulary, as well as to ensure the high-quality encounters with words.

Here, rich instruction plays a particular role. In order to develop an extensive knowledge (vocabulary depth), the learner should encounter the target word in multiple situations over a period of time. The instruction may include both written and spoken forms, collocations, word parts, and meaning. This type of rich instruction helps develop comprehensive knowledge of a word. However, since it is time consuming, it should only apply to the most frequently used words (Webb & Nation, 2017, p. 284). These encounters should also include, in addition to deliberate noticing, repetition, spoken and written retrieval, vocabulary-focused meaning comprehension tasks, dictation, and intensive reading (Webb & Nation, 2017, p. 126).

The abovementioned findings simply indicate the necessary vocabulary size for the language user to communicate freely in each setting. It should be pointed out that the reported numbers do not reflect the learning burden of a single word family. The assumption is that if a learner knows a given word family, the learner knows all of the word's lexical forms. Schmitt (2008, p. 332) points out that to learn such an enormous variety is one of the most challenging aspects of learning English as an L2. Nation (2006, p. 66-67) exemplifies what it means to know a word family with the help of the word *nation*. The word family includes the following words: *national, nationally, nationalizing, nationhood, nationalistic, internationalists, nationalistically, nationalize, nationalized, nationwide*, etc. Achieving the same level of knowledge as a native speaker would take many years of intensive exposure to language.

There are reasons why the learner's lexical competence differs from that of a native speaker. Already in 1976, Jack Richards pointed out the kinds of vocabulary knowledge that a native speaker may have: the vocabulary increases well into adulthood, and there is an understanding of how frequent words appear in print and in speech, together with the collocations of the word. Moreover, native speakers understand the register variations, such as situational constraints or social roles, they know the behavior of the word in a sentence, the word's derivations, its root, and possible associations with words from other languages. Lastly, he or she will be able to name various meanings of the word (polysemous words) and the semantic value (its meaning according to the interpretation – from Schmitt, 1997, p. 8).

These qualities are not necessarily typical for a learner of English as L2. Consequently, the frequency lists should also ideally make a distinction between native speakers and language learners. The lexical challenges compared between L1 and L2 learners are complex. For

instance, the kind of input one is exposed to varies, depending on whether it is L1 or L2 input. According to the Incidental Learning Hypothesis, it is the volumes of reading, and interaction with written language, that seem to influence vocabulary growth to a large extent (Webb & Nation, 2017, p. 50). Another issue is the quality of the input, encounters in different situations, and the time used for processing (Webb & Nation, 2017, p. 57). It is the incidental vocabulary learning that accounts for a large amount of vocabulary acquisition. Conversely, lack of the abovementioned factors may account for the fact that very few L2 learners manage to learn the 9,000 most frequent words (Webb & Nation, 2017, p. 45)

The following table shows the vocabulary knowledge of learners of English as a second language in various countries. The measures investigated the vocabulary size by testing the receptive skills only, shown in the table, alongside the number of hours of instruction received:

Table 4: The receptive vocabulary size of learners of English from various countries, and the amount of instruction they received, based on Laufer (2000).

Country	Vocabulary size	Hours of instruction
Japan EFL University	2,000	800-1200
China English majors	4,000	1800-2400
Indonesia EFL University	1,220	900
Oman EFL University	2,000	1350+
Israel High School Graduates	3,500	1500
France High School	1,000	400
Greece High School	1,680	660
Germany High School	1,200	400

As it appears, the learners that were tested for their receptive vocabulary sizes fall short of meeting the requirements for successfully engaging in communicative activities listed in Table 1. Even at university level, the highest vocabulary size is 4,000 word families, achieved by Chinese learners majoring in English. One can assume that the vocabulary learning techniques were ineffective, and perhaps the emphasis was on grammatical correctness rather than communicative skills. As Laufer remarks, if the students learned 2-3 words per class, the instruction contributed to the students' understanding in a very limited manner, and it is safe to assume that the taught words were actually targeted during the class, rather than randomly picked up by the students during reading (Laufer, 2000, p. 48). Moreover, Schmitt mentions the students' willingness to achieve better vocabulary knowledge as one of the factors facilitating

learning and, as a result, better comprehension and communicative skills (Schmitt, 2008, p. 333).

In the Hungarian perspective, Orosz (2009) reports that the primary school students have the capability of learning approximately 300-400 lemmas in English in a time span of a year (Orosz, 2009, p. 185). A mean vocabulary size estimate for a third-grader was 348 words, for 4th grade: 696, for 5th grade: 1177 and for the 6th grade: 1457 (Orosz, 2009, p. 190). Even though the results showed a great variation among students, some children were able to learn up to 1,000 lemmas during a school year (Orosz, 2009, p. 185). Another finding suggested that deliberate vocabulary teaching had a close connection with the students' vocabulary acquisition and retention (Orosz, 2009, p. 192).

Another insight into learners' vocabulary size was proposed by Schmitt (2010). In his study, the undergraduate students obtained a receptive vocabulary size of 5,000 word families (Schmitt, 2010, p. 199). On the other hand, PhD students were able to score as many as 9,000 word families in a receptive vocabulary size test (Schmitt, 2010, p. 199). This shows a large disparity with the vocabulary sizes of the students in the table above.

A comparative analysis in the Scandinavian perspective is perhaps more relevant in this context, regarding the receptive and productive vocabulary size. Due to the similar linguistical background and similarity of attitudes and approaches towards language learning, the learners across Scandinavia may have a similar point of departure when it comes to learning English vocabulary. Henriksen and Danelund (2015) point out discouraging evidence for the receptive and productive vocabulary size of Danish learners. The targeted group (Danish learners in their first year of upper secondary school) has shown very limited receptive vocabulary knowledge, with more than 80% of learners not having reached receptive mastery of the first 2,000 most frequent word families (Henriksen & Danelund, 2015, p. 36). In their written production, the students relied mostly on K1 frequency level words. Vocabulary employed from AWL was very limited, the number of lexical errors was vast, and the lexical richness measures (such as TTR or LD) were minimal (Henriksen & Danelund, 2015, p. 36).

3. Materials and methodology

3.1 Introduction

In the following, the process that led to collecting the data, the data collection and the following data processing will be described. The data material consists of the results from the students' vocabulary size tests conducted using the free online software (www.my.vocabularysize.com), which is described in more detail in section 3.3. This part of the research includes quantitative analysis. In addition, the participants' essays have been collected to investigate their productive vocabulary size, as well as other lexical measurements. The free software that has been used for that purpose (to be found on <https://www.lextutor.ca/>) is more thoroughly described in section 3.4. This part of the data collection entails qualitative analysis of the learners' productive vocabulary, using measurements that help describe and break down their knowledge on a more individual level. This part of research includes some quantitative elements, as some of the measurements may be quantified.

The two methods are used in the strong conviction that combining qualitative and quantitative methods is the best approach to gaining a broader picture of the students' lexical proficiency. The quantitative approach allows the use of absolute numbers and to be able to generalize information from them, such as 'What is an average productive vocabulary size of a 15-year-old pupil at a Norwegian upper secondary school?', 'How big is the gap between the student who knows most and least word families?' or finally, 'Do the students know enough vocabulary to follow daily conversation or read a novel in English?'. In contrast, the qualitative approach may help investigating the individual vocabulary profiles of the participants, point out their usual errors, the collocations that are used, and establish their lexical sophistication or density. None of the methods seem to be the preferred one. However, together they help paint a broader picture of the proficiency of the students.

The following section sheds light on the informants that participated in the study, as well as the data material that will be used. Both software pages used for the data collection are described, together with the advantages and drawbacks to their usage. More detailed motivation for choosing these two tools is presented, together with the possible limitations. Moreover, the ethical issues and challenges connected to reliability and validity of the study will be discussed.

3.2 Informants and data material

The data material for this study has been collected at Vg1 (first level of upper secondary school) and consists of test of the learners' receptive vocabulary size and a test the students handed in as a part of their ordinary course work. The tests written by the students concerned the assessment of their productive vocabulary. The student group consisted of 100 students, divided into four different classes. The students were enrolled in an English course with classes two-three times a week, a total of five school lessons (45 minutes) each. The students were 16 years old, with one exception where the student was 17. The majority of the students had Norwegian as their mother tongue, while some students had other languages as their first language (one Polish student and one Arabic student). English was the second language for all the students, except for two, and they had been learning the language for 11 years. The observation was pre-approved by the head of the school, as well as written consent from the students being obtained.

Prior to the data collection, the participants were given extensive information about the study. They were presented with the background and the purpose of the study: it has its source in a master's thesis in Literacy Studies, where the purpose was to gather information about the receptive and productive vocabulary size of the students enrolled in a Norwegian upper secondary school. The data collection was an anonymous process, where the students were codified only in order to match their scores from the vocabulary size test, and their performance regarding the productive vocabulary size. This was done so that one can still attempt at making inferences between the two proficiencies, while ensuring the participants' anonymity.

Consequently, the participants were presented with a consent form, so it is easier to ensure that they had received the information. An example form is enclosed in the study and was written in Norwegian to prevent possible misunderstandings (see Appendix 2). The information has also been communicated in a spoken manner to the students, who could have asked questions during that session. Later, when conducting the vocabulary size test, the learners were asked not to be in contact with each other.

As already mentioned in section 2.4, children growing up in Norway are exposed to a large amount of English-speaking media on a daily basis. When enrolling on an English course as a part of their education, learners have already had their first experiences with the language, either in their home country or abroad. As it is not the focus of this study to determine the

influence of either the classroom activities or the extramural impact on their proficiency, no further distinction has been made as to the source of the students' vocabulary size.

The data material then obtained from the VST and VP was carefully analyzed. The scores of the participants' receptive vocabulary sizes were analyzed in the different groups, as well as together. Minimum, maximum, average and median scores were calculated for each group and the whole population. Moreover, the mean decision time for each participant has been included. When it comes to the qualitative part of the study, the essays of the high-, average- and low-scoring students were analyzed. The essays were shortened to count 200 words each before they were submitted to the software. Research suggests that the text length has an impact on measures of lexical richness. Especially the type-token ratio has been shown to be affected by the number of tokens. For stability, it has been suggested that texts of at least 200 tokens provide satisfying stability (Schmitt, 2000, p. 75; Koizumi, 2012, p. 67; Laufer & Nation, 1995, p. 314; McCarthy & Jarvis, 2010, p. 386). Statistical stability is ensured because adding length to the texts would not affect the measurements significantly. Thus, having submitted 200-word texts was enough to obtain a picture of the lexical measurements. The results are presented in detail in Chapter 4.

The targeted population was Vg1 upper secondary school students at one school due to their availability. The sample included 100 students from four different classes in the same school. It was a convenience sample, as the students were willing to participate and were enrolled in the given institution the author had access to. According to Dornyei (2011, p. 100), a rule of a thumb suggests that the sample should ideally involve at least 100 participants. However, there are exceptions concerning the sample size: if the study includes correlational differences, at least 30 participants should be included (Dornyei, 2011, p. 99) Theoretically, for a quantitative study, the sample may be infinitely large. In general, the sampling issue is inherently a practical question, as there is no possibility of including every member of a population. The number of participants should still be large enough to even out the individual differences (Schmitt, 2010, p. 150). With 100 students voluntarily participating in the research, with arbitrary number of boys and girls, of various backgrounds and interests, one may hope that the sample is to some degree a fair representation of the population.

At the same time, the representativeness is compromised, as the sample has a fairly homogenous composition: the students participating in the research were enrolled in general studies. Thus, the vocabulary size of students enrolled in other programs, such as vocational students, was not investigated. According to Dornyei (2011, p. 99) the convenience sampling

and to what extent the results from such sampling may be generalized, is marginal. The sample size allowed for discovering patterns and statistical analysis, as well as an in-depth analysis of the material. Yet, having described the limitations of the sampling above, the practical aspects of such sampling may be assumed to be reasonable for a master's thesis and still allow for valid statistical considerations. No general claims can be made from the results to concern the whole population.

Concerning the qualitative part of the study, extreme case sampling was devised, which entails picking the extreme cases of the research. (Dornyei, 2011, p. 108). The most extreme cases from the sample were chosen on the basis of their receptive vocabulary size score: high-, average- and low-scoring students. This measure was undertaken to check both for common core elements ('What academic vocabulary is used both by highest- and lowest-scoring students?'), but also for the limits of experience ('How does the lexical density of the highest-scoring student differ of that of the lowest-scoring one?'). This type of sampling hopefully increases the scope of the analysis (Dornyei, 2011, p. 129). Another important aspect is lowering the time and effort spent on analyzing the texts, without compromising the validity of the results.

3.3 English Vocabulary Size Test Classic Edition

To measure the receptive vocabulary size of the learners, a vocabulary size test available at my.vocabularysize.com was used. The test was based on Paul Nation's lifelong contributions to research within second language vocabulary learning. It is built on the assumption that high-frequency words are more likely to occur in communication and easier to understand than low-frequency words. This is a version with 144 multiple choice style question. Ten items come from various 1000-word family levels. Thus, the test includes word families until the 25th frequency level. The monolingual version of the test, applied in the current study, was based on the frequency word lists derived from the British National Corpus (BNC – Nguyen & Nation, 2011, p. 87). The test measures vocabulary size quantitatively. The following figure shows the layout of the VST:

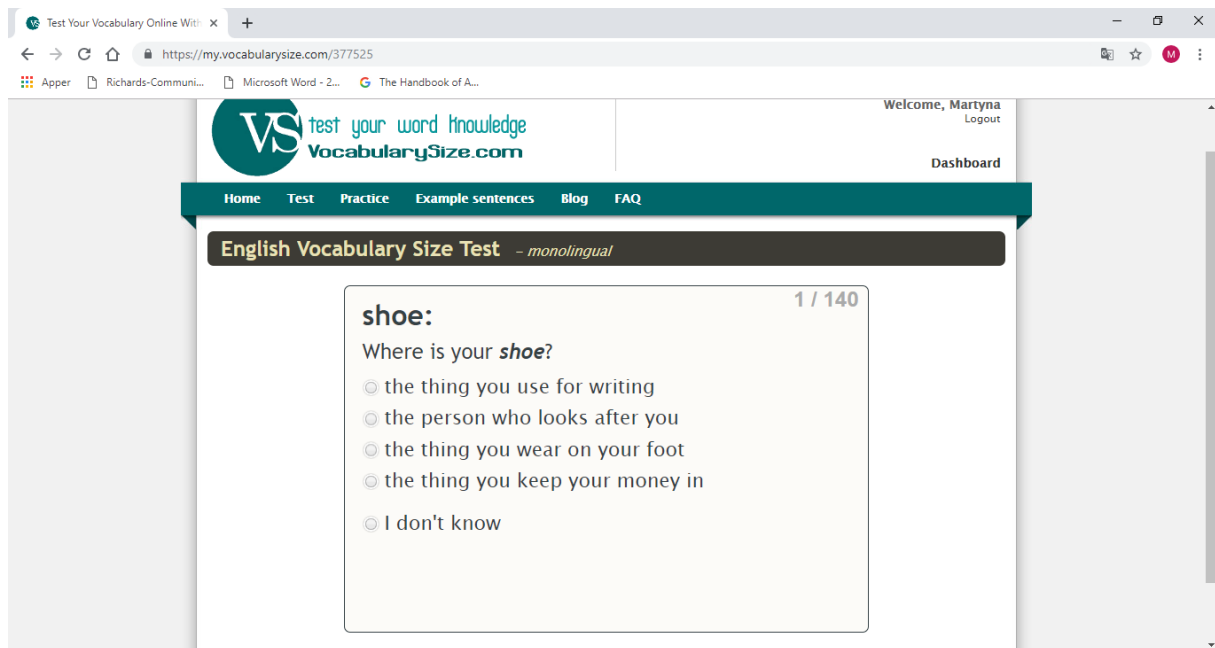


Figure 1: The layout of the Vocabulary Size Test used in the study.

The key concern is to establish reliability, meaning that the results should be replicable, and validity, which ensures that the test measures exactly what it is meant to measure and that the results do not have any other causal relationships than what is being assumed (Bryman, 2001 p. 80). These features may be ensured by the design of the research.

A vocabulary size test may help gauge the growth of vocabulary size of second language learners and help compare it with for instance growth of native-speakers (Nguyen & Nation, 2011, p. 87). A VST may be a useful diagnostic tool for gauging students who may struggle with comprehension of material presented in classes. In addition, as shown in table 3 in section 2.17, a learner needs a target receptive vocabulary size of approximately 9,000 to be able to participate in range of activities involving spoken and written texts in English. Unless the teacher is familiar with the vocabulary size of the students, the vocabulary component of language teaching may be flawed (Nguyen & Nation, 2011, p. 87).

A number of issues need addressing when measuring receptive vocabulary size. The standard, computerized format allows for simplicity and time-saving aspects of the test. Moreover, the test is designed to account for word families, which helps avoiding over-counting. Also, word families are categorized due to their frequency, then grouped into a representative list of word families from the given frequency level. Word families are an appropriate and valid unit of measure when estimating the vocabulary size, thus making the online test a viable alternative and reliable measuring technique. As mentioned above, the

concept of word families is an established unit of measure in linguistic research (Read, 2000 p.19; Schmitt, 2000 p.148).

There are some aspects of the test that may influence the results. Some of the students may be guessing. With four alternatives, the chance of guessing the correct meaning is 25%. The test does not allow for adjusting to guessing (Nguyen & Nation, 2011, p. 98). Moreover, the British National Corpus (BNC), which served as a backdrop for devising the test, may be seen as a poor reference list for Norwegian teenagers. In addition, it is important that the sentences in which the target word occurs, does not give away the answer. This would misrepresent the results (Nguyen & Nation, 2011, p. 98). However, the test has been shown to work well and is recommended for use to gauge the students' receptive vocabulary size (Nguyen & Nation, 2011, p. 99; Nation & Beglar, 2007, p. 9).

3.4 Vocab Profiler (VP), Classic Edition

When it comes to productive vocabulary, the students submitted an essay in the middle of the autumn term. The essays were then pasted into and processed by a software called Vocab Profiler, Classic Edition, available on-line (<https://www.lex tutor.ca/>). The program, developed to perform lexical analysis, conducted an analysis of the texts concerning their vocabulary profile. The program was an adaptation of a lexical tool originally designed by Batia Laufer and Paul Nation, later adapted by the Montreal-based linguist Tom Cobb. Relatively simple, the software helps categorize the vocabulary used, and thus gauge performance of the language learner (Cobb, 2002). The analysis included the word families, types and tokens categorized into the levels of frequency. Additionally, the usage of words mentioned in AWL was recorded. The texts were analyzed according to the use of words from various frequency levels. The natural assumption was that the use of words from the lower frequency levels decreases. Moreover, a more detailed, qualitative study of the students' written production was conducted.

The VP software analyzed the participants' essays. Since the output is sensitive to text length, the essays were shortened to the first 200 tokens in each analysed essay (Schmitt, 2010, p. 252). The output is organized in 25 frequency levels. The starting point is the 1,000 most frequent words (K-1) up to the 25,000 most frequent words in English language (K-25). The program is comprised of the 25,000 words which are integrated from the earlier prepared lists of the (BNC) and the Corpus of Contemporary American English (COCA). The advantage to

using the VP is the practical aspect: it only takes a few seconds of submitting the input to obtain an extensive analysis. It covers word families up to K25 frequency level, beyond which it is unlikely that the students would score in their productive vocabulary. In addition, the software measures the coverage of words from AWL, allowing for analysis of the academic vocabulary.

In this part of the research, a qualitative approach was used with specific emphasis on interpretation of the essays. The written production of language learners has always been a significant part of assessing linguistic proficiency (Gregori-Signes & Clavel-Arroitia, 2015, p. 348). Knowledge about students' lexical richness can be evidence of progress and may help teachers provide appropriate materials and reflect on their strategies. This included the individual texts and what vocabulary had been used, the lexical richness, which entails lexical sophistication, density and variation, use of collocations and lexical errors. This method was applied in order to investigate the learners' vocabulary more individually (Bryman, 2001 p.20). The focus was to describe and analyze, then perhaps suggests some theories regarding pupils' current vocabulary. It is worth mentioning that VP provided quantitative measurements.

As a scientific approach, qualitative research is not applicable as a set of linear steps. It treats theory as something that emerges out of a collection of data, rather than testing theories specified in advance of data collection. Some researchers argue that qualitative research can and should play an important role in relation to quantitative methods and the testing of the theories (Bryman, 2001 p. 268). Thus, this thesis included both types of research. In this way, through the quantitative method's rigorous, precise, controllable and generalizable nature, and through the qualitative method's exploratory, interpretive, and flexible data material, the study may provide the readers with convincing evidence of the state of vocabulary knowledge among the researched group (Dornyei, 2011, p. 35-40). The two methods complement each other and help in gaining a broader, more nuanced picture. Notwithstanding the differences of the two approaches, both adapt the concepts of internal and external validity and reliability. This is due to the fact that incorporating multiple research methods increases the transferability of the findings outside the observed evidence (Dornyei, 2011, p. 46).

An alternative measurement unit is sometimes referred to as a type-token ratio, as it is obtained by dividing the number of different word types by the total number of tokens in the text (Schmitt, 2000 p.74). Using this unit of measurement for lexical variation has its weaknesses, as an article in a scientific journal, for example, will probably have many repetitions, suggesting that its lexical density is low, which is not necessarily the case (Schmitt,

2000, p. 75). The simplicity of both of these lexical measurements does not withstand the fact that they may be used to indicate the overall difficulty level of the text corpus.

The VP has been proven to provide stable measurements of students' writing production. It is almost completely computerized, the results are replicable, and two students of different linguistic knowledge can be distinguished (Laufer, 1995, p. 319). Direct focus on vocabulary is its major strength. Various steps had to be taken into account when using the VP. There is evidence that the software gives less stable results with a text shorter than 200 words (Levitzky-Aviad & Laufer, 2013, p. 134). The corpus had to be checked for whether any essay was shorter than the required length. Moreover, as the type-token ratio is sensitive to the length of the text, and as it is one of the measurements of lexical variation, the submitted essays were shortened to 200 words each. This is the case for measurement of every feature analyzed in the study.

3.4.1 Lexical variation

Lexical variation is one of the measures that helps estimate a learner's lexical richness. It is reasonable to expect that more advanced writers will use a variety of different words rather than repeatedly use the same range of expressions (Read, 2000, p. 200). This variation can be obtained by using synonyms, precise vocabulary, and avoiding repetitions. If most of the words are used repeatedly, then fewer words (types) are familiar to the student. A commonly used measure to estimate a learner's lexical variation is the type-token ratio (TTR). The TTR is measured by dividing the number of different types by the total number of words in the text (tokens - Schmitt, 2000, p. 74). A low TTR may suggest a small vocabulary size, repetitive language or sticking to a subject, thus using similar words to express an idea (Covington, 2008, p. 1). A low TTR means that the writer knows few types, however, it does not need to be a negative factor. In a communication situation, a text with few types (low TTR) will be easily understood by a person who learns the language and who perhaps has not achieved high proficiency yet.

Research reports a range of 0,36 to 0,57 when it comes to the TTR of written texts (Schmitt, 2000, p. 75). An average TTR would be 0,40. The lower the ratio, the more repetitions are to be found in the text. This ratio is greatly affected by the length of the text. This is due to the fact that as the text gets longer, the number of new types does not get relatively bigger

(Schmitt, 2000, p. 75). The central words are rather used repeatedly. The higher the ratio, the less repetitions there are and the more complex the lexical treatment. Despite its flaws, such as high sensitivity to text length, TTR is a widely used statistic measure, which may be computed quickly (Covington, 2008, p. 1).

3.4.2 Lexical sophistication

Another aspect of range of expression is lexical sophistication (LS). The use of technical terms, jargon and the selection of other low-frequency words that are specific to the topic and style of writing are measured and described as a part of LS (Read, 2000, p. 200). Vocabulary of a general nature, here represented by the GSL, is not taken into account. Sophistication of a learner's productive vocabulary size may also be described as 'rareness', or the number of words that the students use that is not expected at that level of education from a learner of English as L2 (upper secondary school). Alternatively, it may be defined as the words that students are not expected to know. There are no guidelines available as to what words are those, or beyond which frequency level it is. It may be measured by estimating the number of sophisticated word families in the text and dividing it by the total number of word families in the corpus (Read, 2000, p. 203).

LS may be considered the most subjective measurement of lexical richness. Even though it is calculated using a general equation, the constituents of that equation are only vaguely defined. What words are considered 'advanced' for the learner should be individually assessed regarding the student's level (Laufer & Nation, 1995, p. 309). Consequently, a given corpus may be analyzed from various perspectives, depending on how 'advanced vocabulary' is defined. Such a vague definition makes it an unstable measure. Moreover, it makes it inapplicable to larger datasets. For a more comparative dataset and conclusions, one should consider other measurements of lexical richness (Laufer & Nation, 1995, p. 310).

Regarding the word-frequency tables, and the expected rareness of the words used by students, one may make certain assumptions. The students were presented to tasks, and thus the written production was not entirely free. Two texts that are about the same subject and involve similar arguments and ideas, are likely to include words of similar frequency of occurrence (Covington, 2008, p. 4). To some extent, the measurements will be similar for the corpus that include texts on the same topics. Greater variation would have been obtained, if the texts were written on a variety of subjects.

For the purpose of this study, students from different classes were targeted. It would seem unreasonable to define sophisticated words randomly, as it may happen that different teachers in those classes target different vocabulary during the course. However, noticeably, the number of tokens for each frequency level decreases in the texts. The higher the frequency level, the fewer tokens are used. The productive vocabulary size has been estimated to be on average 5,000 word families (to be discussed in Chapter 4). Thus, the advanced words are the ones from K6 frequency level and above.

3.4.3 Lexical density

The reason why various measurements of lexical sophistication are applied is that not all words in a text are equally weighted. Lexical density is measured by calculating the percentage of content words in a corpus (Read, 2000 p.196). This was calculated by dividing the total number of content words by the total number of tokens in the text. A text is more lexically dense where the proportion of content words is higher (Schmitt, 2000 p.75; Read, 2000, p. 203). The higher the proportion of content words in a corpus, the more lexically dense the text is (Schmitt, 2000, p. 75). An idea may be presented in a very concentrated way (high lexical density), which is often the case with written texts. The situation is the opposite with spoken texts, where lexical density tends to be lower (Read, 2000, p. 200). LD was a variable used to determine a text's difficulty by checking the diversity of the vocabulary used.

3.4.4 Lexical errors

The TTR discussed in section 3.4.1 is one of the measures that may help describe learners' vocabulary knowledge. However, it takes into consideration only the number of words produced by students. The possibility of the words being used wrongly is not accounted for. In many ways, analysis of lexical errors may be understood as the reverse notion of lexical richness (Read, 2000, p. 204). An analysis of lexical errors, as opposed to grammatical ones, imposes two challenges: identification and distinction from non-lexical mistakes. Lexical errors may be roughly divided into word choice and word form errors. The following table, as adapted by Read (2000, p. 206), shows a more detailed classification and indeed a multitude of possible lexical error types:

Table 5: Classification of lexical errors (based on Read, 2000, p. 206).

I. Lexical choice	A. Individual lexical items	1. Semantically unrelated	a. It has some <i>meanings</i> to study here.
			b. We can help ourselves with <i>doing</i> as the others.
		2. Semantically close	a. They have to <i>come back</i> to Rome.
			b. We can study some <i>strange</i> subjects.
	B. Combinations	1. Two lexical items	a. Young people can <i>say</i> their <i>ideas</i> .
			b. I will <i>bring to go</i> the development in the area.
		2. Phrases	a. We have a <i>lot common</i> .
			b. You can everything from the <i>bottom of your mind</i> .
3. Multiple errors (core lexical item)		a. It is being popular <i>year and year</i> .	
		b. I will get English comprehension perfectly.	
II. Lexical form	1. Derivational errors	a. Keep the class more <i>activity</i> .	
		b. There are few <i>conflictions</i> here.	
		2. Verb forms	a. I <i>say</i> it to you yesterday.
			b. She <i>like</i> the story.
	3. Phonetically similar – Semantically unrelated	a. She <i>things</i> it may help her.	
		b. The students need a <i>wild</i> horizon.	
	4. Major spelling errors	a. It was impossible to <i>akeiwe</i> that dream. (achieve)	
		b. The line was drawn <i>heriosontaly</i> . (horizontally)	

The percentage of error occurrence in a text has been researched to some extent (Read, 2000, p. 202). The investigation of lexical errors is an important part of lexical analysis, since writing of second language learners is characteristic in that it often includes wrong form choices and forms. A native speaker's writing is not likely to include such mistakes (Read, 2000, p. 200). Besides, no other measurement takes into account possible wrongful usage of lexical items. Consequently, a learner may use very advanced and varied vocabulary, scoring very high on both LS, LD and LV, but these words may not constitute a meaningful, correct text. One of the weaknesses is that a lexical error analysis does not account for the seriousness of the errors. For instance, minor spelling mistakes are not listed in the table above. The learner may be familiar with the word and use it correctly, even though one letter has been misplaced (*horizontaly* instead of *horizontally*). Ideally, the analysis should include a discussion of whether the mistakes pose difficulty in interpreting the text. If so, they are more serious and lower the quality of the text than, for example, misplacing a letter in a word.

3.4.5 Collocations

As a lexical measurement, collocations impose various statistical difficulties when it comes to measurement (Schmitt, 2010, p. 124-130). Identification makes it tricky to research collocations and chunks of language (Schmitt, 2010, p. 120). No particular measurement tool was used to check the submitted samples of texts for collocations. Instead, the sample texts were manually checked for the occurrence of collocations. The following sources on collocation of native-speakers were consulted: the hundred most frequent collocations in English (Shin & Nation 2012, p. 346-348), the word frequency list in American English (Davis & Gardner 2010), and the List of Frequent Collocations by Maria Moreno Jaen (Jaen 2017). As there are many various kinds of collocations, this study was limited to check for the most common lexical collocations, as presented below:

Table 6: Types of lexical collocations used in the study (based on Gitsaki, 1999, p. 59).

No.	Type	Example
1.	Noun Preposition	argument about
2.	Adjective Preposition	happy about
3.	Verb Noun	make an impression
4.	Adjective Noun	weak coffee
5.	Noun Verb	dogs bark
6.	Noun1 of Noun2	a piece of information
7.	Adverb Adjective	deeply disturbed
8.	Verb Adverb	affect deeply
9.	Preposition Determiner Noun	on the other hand
10.	Phrasal Verb	to pass out

3.5 Ethical issues

The material collected as a part of this study may seem unproblematic in terms of sensitive, personal information and privacy. However, all research should be carried out in a way that does not compromise personal rights to remain anonymous, unless otherwise indicated. Ethically performed research should be based on openness, disclose information of the full extent of what participating in the research implies, applying to the appropriate agents for approval of the research and, last but not least, responsible data collection, retention and finally, destruction (Dornyei, 2011, p. 63-69).

According to the University's guidelines, the NSD (Norwegian Centre for Research Data) needs to be notified about the research, the involvement of the participants, their age and the purpose of data collection. Written consent was obtained from the participants, with emphasis on the voluntary nature of their participation, as well as the fact that the students at any point of the testing may have withdrawn their involvement without consequences. It was emphasised that any decision on their part would bear no consequence on the future dealings. Written consent forms were stored safely until the end of the research.

Regarding data collection, the students' results needed to be kept anonymous. At the same time, it was necessary to link the receptive and the productive vocabulary size to the same student for the purposes of the analysis. The class lists were enumerated in no particular order. After the results were matched, any record that would enable tracking the results to an individual were destroyed.

3.6 Reliability and validity issues

A test is valid if it measures what it is assumed to measure (Dornyei, 2011, p. 51). Here, a sample rate may be helpful: should one wish to test knowledge of the most frequent 1,000 words in a language, if the test includes 100 words from that level, then the sample rate is 1 in 10. One may consider this a sufficient sample rate. If this measurement was 1 in 50 (20 words out of a 1,000), then the trustworthiness of the results is seriously compromised (Schmitt, 2000, p. 166). The higher the sample rate, the more valid the vocabulary test.

Read (2000, p. 81) calls into question the validity of tests measuring vocabulary knowledge. For instance, the choice of measurement techniques is not always properly justified. Moreover, there is no certainty that the vocabulary measuring tests do not test other aspects of language proficiency at the same time. In the early stages of vocabulary testing, it was considered an objective measurement if the target words were presented in isolation, for instance as a list, or in a multiple-choice test (Read, 2000, p. 99). A longer, contextual clue was considered to transform the activity into a reading-comprehension test. Nonetheless, the role of context should not be underestimated. A word may have multiple meanings or belong to multiple parts of speech. Although there is hardly any existing research that documents the function of contextual clues in measuring vocabulary knowledge, it has nevertheless been a widely applied approach to provide background for the target words (Read, 2000, p. 101). A

major part of the research conducted in the twentieth century on vocabulary knowledge comes from separate testing independent from the setting (Read, 2000, p. 115).

A wide variety of purposes may be considered in vocabulary testing. The choice of testing software should be solely justified by the purpose the research is to serve. Poor choice when it comes to the measurement may undermine both reliability and validity. The most commonly used is arguably the achievement test, which measures the number of words learned throughout a course. A diagnostic test may be applied to gauge students' level of knowledge, so that specific attention may be devoted to the material that is not so familiar. Alternatively, a placement test may be conducted to facilitate placing students in appropriate classes. The purpose of testing may be of a more general nature. One may try to test the breadth of vocabulary knowledge, that is how many word families the learners know. This is the sole purpose of this study. Another measure that may be estimated is the quality of that knowledge, also known as vocabulary depth (Schmitt, 2000, p. 164). It encompasses how well the students know a given word, for instance by knowing the word's collocations. A student may use quite sophisticated and low-frequency vocabulary in text production, yet not knowing how and when to use those words. Collocational knowledge helps determine how well the word is known to the learner.

As specified, the purpose of the underlying thesis is to measure the learners' total vocabulary size. For that purpose, sample words from various frequencies were needed to be included in the test. The test has been a subject to many revisions (Nizonkiza, 2016, p. 171). The selection of words is guided by the frequency with which they occur, and the current frequency list consists of 25 bands now (Nizonkiza, 2016, p. 172).

When it comes to the reliability of vocabulary testing, the measurement should be retained over time. If a student takes the vocabulary test, the test should produce the same results at any point in time (also referred to as perfect reliability – Schmitt 2000, p. 166). However, the reality is quite different: the student's motivation, interest, fatigue, lack of time, the environment, and the nature of the test will all influence the results. Again, the sample rate may help adjust the reliability to desired level.

Lastly, an issue that may compromise both validity and reliability is the practical aspect of the test. Researchers encounter twofold complications when attempting to sample enough lexical units to ensure reliability and validity of the study, at the same time minimizing the participant's fatigue (Hirsh, 2010, p. 380). Ideally, the student should take a long test, which

tests as many items as possible, in order to give a thorough picture of his or her vocabulary knowledge. To administer a test that lasts many hours will often be impractical, but will produce more reliable results. There is also a danger of students answering less and less carefully the longer the test lasts, again compromising results (Schmitt, 2000, p. 167). As a general rule, all of the three aspects should be weighed against each other for the outcome to be most valid and reliable, but at the same practically achievable.

3.7 Possible limitations of English Vocabulary Size Test Classic Edition

Potential issues regarding using the Vocabulary Size Test (VST), monolingual edition, in the classroom for the purpose of measuring students' vocabulary size will be discussed in the following. Lexical knowledge has been interpreted as a continuum rather than in absolute terms. The level of knowledge starts from a shallow knowledge of the word, through seven levels, until it reaches the final level, where the learner can use the word independently in communication (Laufer and Goldstein, 2004, p. 400). The researcher needs to choose a testing tool so that the results fulfill the criteria for validity, reliability, and practicality.

The VST is a test that tests many items (144) in a multiple-choice format. In some of the cases, the choices are closely related (for instance, where the tested item is *perennial*, the alternative answers are two describing a plant and two describing an animal). Sometimes, the alternatives are of a very different nature and are different parts of speech. Here, the participant draws on partial knowledge. For example, where the word *erythrocytes* is being tested, one only has to know that it has something to do with blood, as the other answers are largely unrelated.

There is a danger that guessing will be a strategy adopted by many of the participants, especially when the test does not allow for skipping the questions or does not allow a blank answer. However, research (Nation, 2013, p. 350) reports that over 50% of answers were chosen due to associations, with a difference between the high-ability group and the low-ability group. The high-ability group guessed 8% of the answers with a 50% success rate, while the low-ability group guessed more (21%) with a lower success rate (35%). This confirms that the participants rather draw on their knowledge and associations than leave the result to accidental guessing, which eliminates one of the major potential problems with VST. Still, one may assume that some of the answers are a result of guessing.

The tests differ with regard to the amount of contextualization, as well as how the items are to be chosen. The design of the test, where a word is given a minimal context in which the definition should be matched from the provided list of possible answers, is obviously only one of the possibilities. The participants may also be given isolated words in a sentence context, or in texts. Matching is found to be easier than supplying item tests, with higher reliability and validity rates (Nation, 2013, p. 353). The context is of major value to the participant: it shows usual ways of use and indicates the part of speech, as well as encouraging normal access to meaning. Providing context thus gives a greater chance for the participant to show their knowledge. The drawback is that it is more time-consuming and fewer items can thus be tested within the given time frame.

The VST is clearly a measurement of breadth of knowledge, which is one of the dimensions of vocabulary knowledge not concerned with context, fluency, or depth (Laufer & Goldstein, 2004, p. 401). Thus, one can only measure how many words are known, without looking into the depth of that knowledge. Some vocabulary aspects will also be investigated from the qualitative perspective, which will help shed light on the depth of the vocabulary. This is the reason behind including the qualitative aspects. Another drawback is that a limited number of words may be tested out of concern for the practicality of the test, and one will thus never truly investigate the whole extent of the participant's lexical knowledge. Other aspects of lexical ability, involving measuring the words that learners do not know well, or to what extent known words are used correctly, are not available for measuring with this test. The most important aspect of word knowledge, the vocabulary depth (quality), is not tested (Gregori-Signes & Clavel-Arroitia, 2015, p. 548). That is why the Vocab Profiler is used to determine other aspects of the participants' vocabulary knowledge.

3.8 Possible limitations of Vocab Profiler, Classic Edition

One possibility for assessing students' productive knowledge is by employing statistical measures which will analyze the texts. Researchers have attempted to meet various challenges regarding lexical production. This includes whether the measures give consistent results (reliability), whether the test's behavior remain consistent over time (Schmitt, 2000, p. 166) or the relationship between the lexical quality and the whole text. Determining the relationship between the productive and receptive vocabulary knowledge of the learners is another important

aspect (Read, 2000, p. 197-198). The available research provides some insights into the challenges encountered when measuring students' productive vocabulary knowledge.

Firstly, Read (2000, p. 199) draws attention to the effect of the writing task itself. The material inserted into the software is the students' answer to (a) task(s) devised by various teachers. The research conducted by Reid in 1990 (Read, 2000, p. 199) reports that the nature, style and content of the task will influence many of the lexical variables. Both the average length of the word, lexical density, as well as pronoun percentage were affected by the kind of task the participants chose. Thus, the appropriateness of the writing task according to the research question should be carefully considered, and any conclusions drawn from the results should be careful interpretations.

Once the students have submitted their texts, one may start to calculate the various statistics. All of them have been identified in section 2.1. To ensure valid and reliable results, consistency is necessary. Also here, many aspects will influence the obtained numbers, and should be carefully interpreted. For instance, text length will influence the type-token ratio (Read, 2000, p. 202). Naturally, the ratio is lower for a 1,000-word text than for a 200-word text, since, as the writer proceeds, fewer new words are used, thus lowering the score. This can be remedied by only taking into consideration the first 200 words of each essay. It is necessary to adjust the length of the essays, especially if it varies greatly across the corpus. One solution is to adjust to the shortest submitted essay (Read, 2000, p. 202).

Another aspect of analysis is whether to count erroneous words (Read, 2000, p. 202). Should the student write a word incorrectly, various reasons may have caused it: hasty spelling, where a typo has occurred (this can be determined if the same word is used more than once, and is written correctly at least once), dyslexia, or unfamiliarity with the word. Regardless of the reason, a choice should be made if the incorrect words should be counted. Then there is the issue of distinguishing between lexical and grammatical mistakes.

Lastly, it should be pointed out that the software does not account for the holistic quality of the participant's performance (Read, 2000, p. 217). Indeed, vocabulary constitutes one of the descriptors of a learner's writing, such as content (relevance, topic knowledge, substantive, well-developed thesis), organization (fluency, logic, cohesion, organization, and clarity of the arguments), language use (construction, subject-verb agreement, number, word, tense, article, pronouns), and mechanics (spelling, punctuation, capitalization, paragraphing – Read, 2000, p. 217). In the most extreme cases, a pupil who knows that vocabulary will be assessed, with

disregard to any other component of language, may simply enter random words, without concern of the syntax, grammar, or coherence of the text. Thus, all the essays should be checked manually and require human involvement, before being subjected to any software.

4. Results

4.1 Introduction

In the following, the results from both the Vocabulary Size Test and Vocab Profiler, Classic Edition, are presented. First, the receptive vocabulary size of each research group (A, B, C and D) are presented. In tables 6, 7, 8 and 9, the receptive vocabulary size score (number of word families) together with the decision time (milliseconds) are presented for each participant. The results are presented in descending order. The number of participants, maximum, minimum, median and average score for each group are also presented, for both the receptive vocabulary size and decision time. Finally, the combined results for all participants across all groups appear in a table.

The second part of the current chapter deals with the qualitative analysis of the productive vocabulary size. The individual text profiles of high-, average- and low-scoring participants from each research group are presented. In group A, two participants scored the same maximum amount of word families: 12,700. Thus, both participants' text profiles were included in the section. For the average score, the participant who scored closest to the average receptive vocabulary size, has been selected. For the low-scoring participants, the student who scored the lowest number of word families in receptive vocabulary size, has been taken into consideration. Thus, in total, 13 individual lexical profiles have been analyzed. In addition to the frequency levels represented in the students' writing, other measurements are compared. This included text coverage with K1 frequency level words, the number of word families and types per frequency level.

Other measurements are listed in the attached tables: TTR, LD, and the number of all types and tokens in the text. These measurements are a part of lexical richness. Additionally, a more qualitative approach has been applied to analyzing the use of academic vocabulary, collocations and lexical errors in the students' writing. These are not estimated by numerical measurements but described in order to obtain a fuller picture of the students' lexical proficiency.

4.2 Receptive vocabulary

The following sections include the comprised results from the VST for all the research groups. Firstly, each research group is shown individually regarding the participants' receptive vocabulary size. The number of word families known by every participant is presented in descending order, together with the mean decision time in milliseconds. In addition, a table with minimum, maximum and average scores for both of these measurements. Finally, a graphical representation of the results is included.

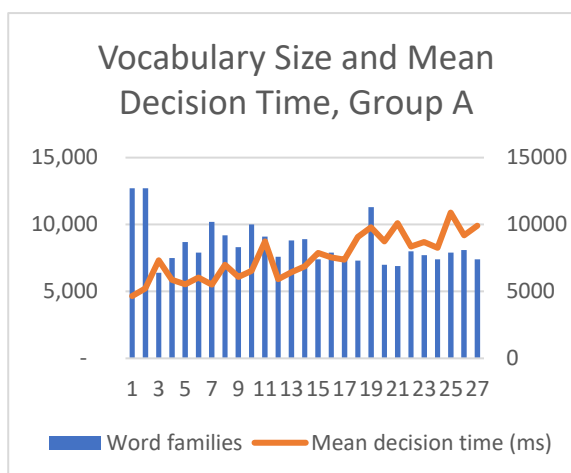
4.2.1 Receptive vocabulary, Group A

Table 7 shows the scores of all of the participants from research group A, as well as the mean decision time, expressed in milliseconds. The highest score in the group was 12,700-word families, with two students receiving this result. These were also the students that submitted the test the fastest. One of the students with this score had the lowest mean decision time (4655 ms). The lowest score was 6,400 word families, with a mean decision time of 7331 ms. The median in this research group was 7,900 word families and 7383 ms. The average scores were respectively 8,515 and 7536,6. Twenty seven participants agreed to be involved in the research in Group A. The results are presented in the table below:

Table 7: The receptive vocabulary size of participants in research Group A.

Participant	Word families	Mean decision time (ms)			
1	12,700	4655			
2	12,700	5217			
19	11,300	9789			
7	10,200	5517	Average	8,515	7536,6
10	10,000	6535	Max	12,700	10891
8	9,200	6984	Min	6,400	4655
11	9,100	8735	Median	7,900	7383
14	8,900	6861	Participants	27	
13	8,800	6449			
5	8,700	5536			
9	8,300	6092			
26	8,100	9184			
22	8,000	8344			
6	7,900	6044			
16	7,900	7543			

25	7,900	10891
23	7,700	8681
12	7,600	5919
17	7,600	7383
4	7,500	5860
15	7,400	7878
24	7,400	8268
27	7,400	9902
18	7,300	9068
20	7,000	8731
21	6,900	10091
3	6,400	7331



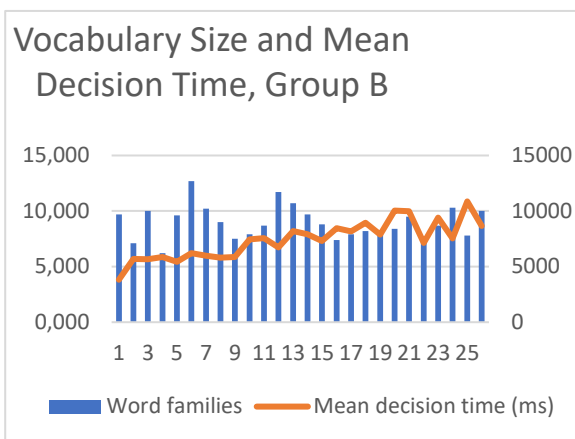
4.2.2 Receptive vocabulary, Group B

Table 8 shows the receptive vocabulary size of participants in Group B, with the scores of the participants and the mean decision time. The highest score in the group was 12,700 word families, with one student receiving this result. However, the student receiving the highest result did not submit the fastest. The lowest mean decision time was 3821 ms, and the highest 10,867 ms. The lowest score was 6,200 word families, with a mean decision time of 5,868 ms. The median in this research group is 8,750 word families and 7485,5 ms. The average scores are 8,973 and 7405,08 respectively. The detailed results are as follows:

Table 8: The receptive vocabulary size of participants in research Group B.

Participant	Word families	Mean decision time (ms)		
6	12,700	6201		
12	11,700	6730		
13	10,700	8188		
24	10,300	7524	average	8,973 7405,08
7	10,200	5978	max	12,700 10867
3	10,000	5656	min	6,200 3821
26	10,000	8654	median	8,750 7485,5
1	9,700	3821	participants	26
14	9,700	7906		
5	9,600	5422		
21	9,500	9971		
8	9,000	5817		
15	8,800	7311		
11	8,700	7561		
23	8,700	9410		
20	8,400	10048		
18	8,200	8948		
10	7,900	7447		
17	7,900	8174		

19	7,900	7886
25	7,800	10867
22	7,700	7132
9	7,500	5852
16	7,400	8456
2	7,100	5704
4	6,200	5868



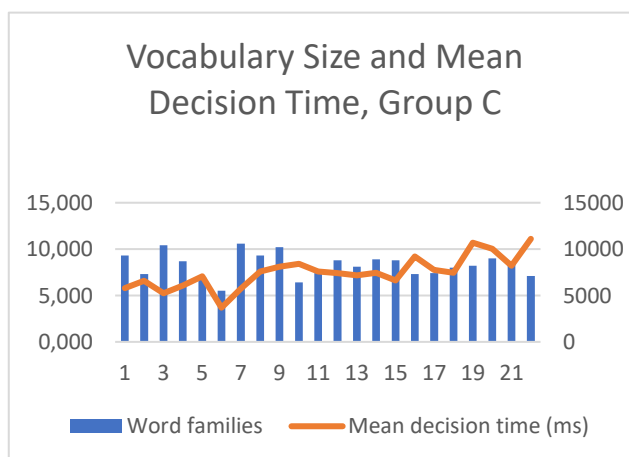
4.2.3 Receptive vocabulary, Group C

Table 9 shows the scores of all of the participants, as well as the mean decision time, for Group C. This was the group with the smallest number of participants, i.e. 22 students. The average score was 8,286 word families and 7,498,3 ms. The median was correspondingly 8,400 and 7,453,5. In this group, the highest score was 10,600 word families. One student scored this high number, with others following with 10,400 and 10,200 word families. Moreover, the highest scoring student submitted the test in time significantly below the average (5718 ms). The lowest score was 5,500 word families and the lowest mean decision time was 3683 ms.

Table 9: The receptive vocabulary size of participants in research Group C.

Participant	Word families	Mean decision time (ms)
7	10,600	5718
3	10,400	5238
9	10,200	8107
1	9,300	5804
8	9,300	7569
20	9,000	10030
14	8,900	7445
12	8,800	7411
15	8,800	6624
4	8,700	6053
21	8,600	8216
19	8,200	10678
13	8,100	7184
18	8,000	7462
11	7,400	7599
17	7,400	7742
2	7,300	6600
16	7,300	9214
22	7,100	11106
5	7,000	7070
10	6,400	8410
6	5,500	3683

average	8,286	7498,3
max	10,600	11106
min	5,500	3683
median	8,400	7453,5
participants	22	

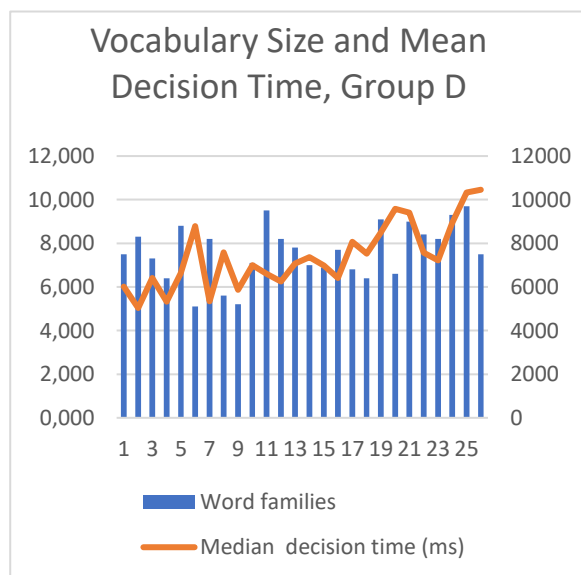


4.2.4 Receptive vocabulary, Group D

Table 10 shows the scores of all of the participants from research group D, as well as the mean decision time, expressed in milliseconds. The highest score in the group was 9,700-word families, with a mean decision time of 10335 ms. The lowest score was 5,100 word families, with a mean decision time of 8790 ms. The median in this research group was 7,700 word families and 7148 ms. The average scores were respectively 7,600 and 7393,8. Twenty six participants agreed to be involved in research in Group D. The results are presented in the table below:

Table 10: The receptive vocabulary size of participants in research Group D.

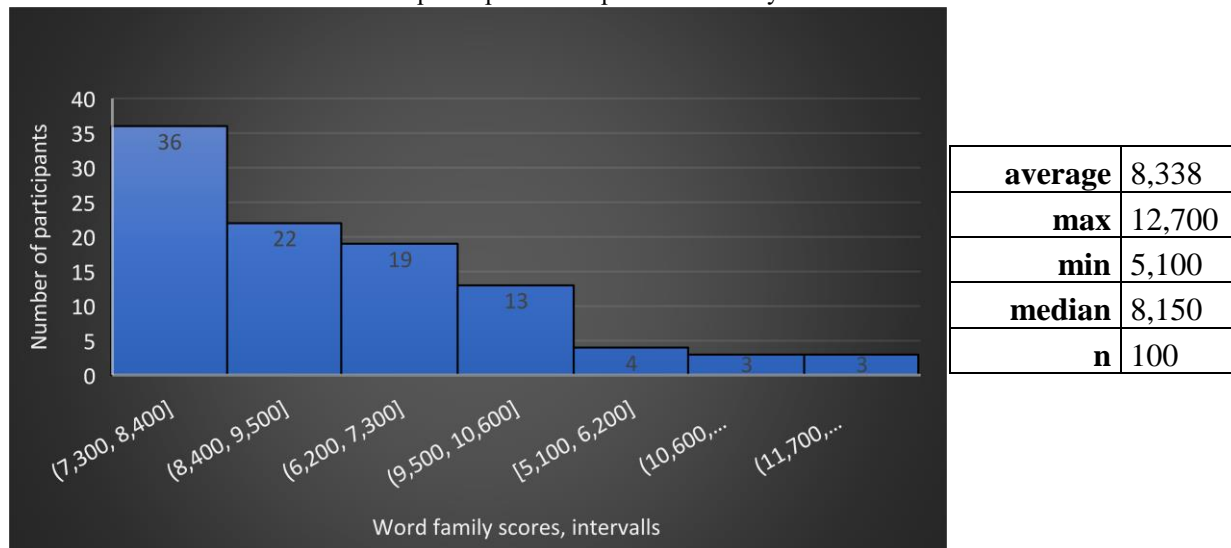
Participant	Word families	Median decision time (ms)			
25	9,700	10335			
11	9,500	6601			
24	9,300	8941			
19	9,100	8490	average	7,600	7393,8
21	9,000	9408	max	9,700	10455
5	8,800	6626	min	5,100	5033
22	8,400	7569	median	7,700	7148
2	8,300	5033	participants	26	
7	8,200	5342			
12	8,200	6246			
23	8,200	7225			
13	7,800	7071			
16	7,700	6406			
1	7,500	6018			
26	7,500	10455			
3	7,300	6404			
10	7,100	6995			
14	7,000	7367			
15	6,900	6983			
17	6,800	8063			
20	6,600	9577			
4	6,400	5320			
18	6,400	7533			
8	5,600	7579			
9	5,200	5864			
6	5,100	8790			



4.2.5 The combined results of receptive vocabulary size

Table 10 above comprises the results of the receptive vocabulary sizes of all of the participants, obtained by the Vocabulary Size Test. Out of the total number of students, the majority (36 participants, or 36%) scored somewhere between the 7,300 to 8,400 word families. The next largest group of students, 22 (22%) in total, mastered 8,400 to 9,500 word families. 19 (19%) of the students knew between 6,200 to 7,300 word families, whereas 13 students recognized between 9,500 to 10,600 word families. The lowest score, between 5,100 to 6,200 was obtained by four (4%) of the participants. Only three participants (3%) obtained the highest score (11,700-12.800) and the next highest score (10,600-11,700). An average participant managed to recognize 8,338 word families.

Table 11: The combined results of all participants' receptive vocabulary sizes.



average	8,338
max	12,700
min	5,100
median	8,150
n	100

4.3 Productive vocabulary profiles, Group A

In the following section, the individual text profiles of high-, average- and low-scoring participants from each research group is presented. In group A, two participants scored the same maximum number of word families: 12,700. Thus, both participants' text profiles are included in the section. For the average score, the participant who scored closest to the average receptive vocabulary size, has been selected. In addition to the frequency levels represented in the students' writing, other measurements were compared. This included text coverage with K1 frequency level words, the number of word families and types per frequency level. Other measurements are listed in additional table, together with tables 10-23: TTR, LD, number of all types and tokens in the text.

4.3.1 High scoring students' vocabulary profiles

Two students achieved the same score in the VST, as described in the previous section. Their productive vocabulary was thus compared. Similarities in terms of use of words from various frequency levels are described, and the kind of collocations used or mistakes made. The first participant, scoring 12,700-word families with a mean decision time of 4655 ms, also agreed to submit his essay as a part of the research. The first student's vocabulary profile shows that his productive vocabulary size was approximately 5,000 word families. The student used 95 words from the K1 frequency level, 16 words from K2, three words from K3 and two from K4 frequency level. There is a gap at K5 level, and two words used from K6 level. A very large coverage (87.4%) of tokens used in the text come from the K1 frequency level

Table 12 The productive vocabulary text profile of one of the highest scoring participants in research Group A.

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. Token (%)
K-1 Words :	95 (80.5)	104 (79.39)	180 (87.4)	87.4
K-2 Words :	16 (13.6)	18 (13.74)	18 (8.7)	96.1
K-3 Words :	3 (2.5)	3 (2.29)	3 (1.5)	97.6
K-4 Words :	2 (1.7)	2 (1.53)	2 (1.0)	98.6
K-5 Words :				
K-6 Words :	2 (1.7)	2 (1.53)	2 (1.0)	99.6
K-7 Words :				
K-8 Words :				
K-9 Words :				
K-10 Words :				
K-11 Words :				
K-12 Words :				
K-13 Words :				
K-14 Words :				
K-15 Words :				
K-16 Words :				
K-17 Words :				
K-18 Words :				
K-19 Words :				
K-20 Words :				
K-21 Words :				
K-22 Words :				
K-23 Words :				
K-24 Words :				
K-25 Words :				
Off-List:		??	1 (0.76)	1 (0.49)
Total (unrounded)	118+?	131 (100)	206 (100)	≈100.00

RELATED RATIOS & INDICES	
Pertaining to whole text	
Words in text (tokens):	206
Different words (types):	131
Type-token ratio:	0.64
Tokens per type:	1.57
Lexical density (content [105]/total [206]):	0.51
Pertaining to onlist only	
Tokens:	205
Types:	130
Families:	118
Tokens per Family :	1.74
Types per Family :	1.10

The second student's individual vocabulary profile is represented in Table 13. Just like the pupil whose vocabulary skills were described above, the score was 12,700-word families, with 5217 ms in mean decision time. There are reasons to believe that the individual lexical skills may have similar characteristics. Here, the pupil used words from frequency levels from 1 through 7, then from the 11K frequency level. Only one word from the lowest frequency level was used ('suffix'). Although the words were used correctly, they may be a result of the nature of the task and may have been included in the task or otherwise included (for instance in the attachment to the task). It is thus difficult to assume that the student would actually be able to recall the terms independently. The productive vocabulary size is estimated to be 4,000 word families.

Other ratios measured by the software, such as the number of tokens and types in the whole text, were similar in both cases (respectively 615 and 677, and 284 and 278). The lexical density (measured by dividing the content words by the total number of words) was 0.51 for the previous participant, and 0.53 for the current one. The detailed results are presented in Table 13 below.

Table 13: The productive vocabulary text profile of one of the highest scoring participants in research Group A.

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. token (%)
K-1 Words :	89 (80.9)	97 (76.38)	176 (84.6)	84.6
K-2 Words :	13 (11.8)	14 (11.02)	15 (7.2)	91.8
K-3 Words :	4 (3.6)	4 (3.15)	5 (2.4)	94.2
K-4 Words :				
K-5 Words :	2 (1.8)	2 (1.57)	2 (1.0)	95.2
K-6 Words :	1 (0.9)	1 (0.79)	1 (0.5)	95.7
K-7 Words :				
K-8 Words :				
K-9 Words :				
K-10 Words :				
K-11 Words :	1 (0.9)	1 (0.79)	1 (0.5)	96.2
K-12 Words :				
K-13 Words :				
K-14 Words :				
K-15 Words :				
K-16 Words :				
K-17 Words :				
K-18 Words :				
K-19 Words :				
K-20 Words :				
K-21 Words :				
K-22 Words :				
K-23 Words :				
K-24 Words :				
K-25 Words :				
Off-List:		??	7 (5.51)	8 (3.85)
Total (unrounded)	110+?	127 (100)	208 (100)	≈100.00

RELATED RATIOS & INDICES	
Pertaining to whole text	
Words in text (tokens):	208
Different words (types):	127
Type-token ratio:	0.61
Tokens per type:	01.64
Lexical density	0.53
Pertaining to onlist only	
Tokens:	200
Types:	120
Families:	110
Tokens per Family :	01.82
Types per Family :	01.09

4.3.2 Average scoring student's vocabulary profile

The average receptive vocabulary size of research group A was 8,286 word families. Thus, a student who scored closely to the average (vocabulary size of 8,200 word families) will be analyzed here. The participant used words from the first four frequency levels, in addition to using one word from the sixth level. The number of different types was 111 and different tokens was 197, with a type-token ratio of 0.39. Lexical density, on the other hand, was 0.56. The receptive vocabulary size is estimated to be 4,000 word families, with 88.4% word families from the first 1,000 words, then 5 words in both K2 and K3 level. There was a gap in both K4 and K5, and there was one word used from K7 frequency band. It is thus difficult to assume that the student was particularly familiar with words from that frequency band.

Table 14: The productive vocabulary text profile of the average scoring participant in research Group A.

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. Token (%)
K-1 Words :	84 (88.4)	97 (86.61)	180 (89.1)	89.1
K-2 Words :	5 (5.3)	6 (5.36)	9 (4.5)	93.6
K-3 Words :	5 (5.3)	6 (5.36)	7 (3.5)	97.1
K-4 Words :				
K-5 Words :				
K-6 Words :	1 (1.1)	1 (0.89)	1 (0.5)	97.6
K-7 Words :				
K-8 Words :				
K-9 Words :				
K-10 Words :				
K-11 Words :				
K-12 Words :				
K-13 Words :				
K-14 Words :				
K-15 Words :				
K-16 Words :				
K-17 Words :				
K-18 Words :				
K-19 Words :				
K-20 Words :				
K-21 Words :				
K-22 Words :				
K-23 Words :				
K-24 Words :				
K-25 Words :				
Off-List:		??	1 (0.89)	5 (2.48)
Total (unrounded)	95+?	112 (100)	202 (100)	≈100.00

RELATED RATIOS & INDICES	
Pertaining to whole text	
Words in text (tokens):	202
Different words (types):	112
Type-token ratio:	0.55
Tokens per type:	01.80
Lexical density	0.50
Pertaining to onlist only	
Tokens:	197
Types:	111
Families:	95
Tokens per Family :	02.07
Types per Family :	01.17

4.3.3 Low scoring student's vocabulary profile

In Table 15, the results of the low-scoring participant in Group A are summarized. In the submitted essay, words from frequency levels K1 through K6 were used. In addition, one token from level K15 and K21 was used. The highest percentage of tokens belongs to the first level of frequency (83.5%). The type-token ratio was estimated at 0.53, with lexical density of 0.56. The productive vocabulary size of that student may be estimated to be 5,000 word families.

Table 15: The productive vocabulary text profile of the lowest scoring participant in research Group A.

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. token (%)
				(%)
K-1 Words :	73 (76.8)	79 (74.53)	167 (83.5)	83.5
K-2 Words :	9 (9.5)	10 (9.43)	14 (7.0)	90.5
K-3 Words :	8 (8.4)	8 (7.55)	11 (5.5)	96.0
K-4 Words :	2 (2.1)	2 (1.89)	2 (1.0)	97.0
K-5 Words :				
K-6 Words :	1 (1.1)	1 (0.94)	1 (0.5)	97.5
K-7 Words :				
K-8 Words :				
K-9 Words :				
K-10 Words :				
K-11 Words :				
K-12 Words :				
K-13 Words :				
K-14 Words :				
K-15 Words :	1 (1.1)	1 (0.94)	1 (0.5)	98.0
K-16 Words :				
K-17 Words :				
K-18 Words :				
K-19 Words :				
K-20 Words :				
K-21 Words :	1 (1.1)	1 (0.94)	1 (0.5)	98.5
K-22 Words :				
K-23 Words :				
K-24 Words :				
K-25 Words :				
Off-List:		??	3 (2.83)	3 (1.50)
Total (unrounded)	95+?	106 (100)	200 (100)	≈100.00

RELATED RATIOS & INDICES	
Pertaining to whole text	
Words in text (tokens):	200
Different words (types):	106
Type-token ratio:	0.53
Tokens per type:	01.89
Lexical density	0.56
Tokens:	197
Types:	103
Families:	95
Tokens per Family :	02.07
Types per Family :	01.08

4.4 Productive vocabulary profiles, Group B

In the following sections (4.3.1, 4.3.2 and 4.3.3), the high-, average- and low-scoring students' lexical profiles from group B are presented. The high-scoring student in this research group obtained a receptive vocabulary size of 12,700, the average-scoring student scored 9,000 word families, and the low-scoring student in this group scored 6,200 word families on the VST. In the following, their individual lexical profiles are presented.

4.4.1 High scoring student's vocabulary profile, Group B

The highest scoring participant in Group B submitted an essay where word families from frequency levels K1 through K8 were represented. One word from the K15 level was used (*kilter*). In level K1, the number of types used was 163 (82.3%). However, already in level K2 and K3, that number became markedly smaller (13, or 6.6% and 10, or 5.1% respectively). For the next levels (K4, K5 and K6), the number of tokens decreased further (4, 1 and 1 tokens respectively). The final two levels (K8 and K15) are represented with one single token each (*doughnut* and *kilter*, as mentioned above). Moreover, the type-token ratio was 0.61, whereas lexical density was 0.55. The productive vocabulary size of the participant was estimated to be 7,000 word families.

Table 16: The productive vocabulary text profile of the highest scoring participant in research Group B.

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. Token (%)
K-1 Words :	79 (73.1)	87 (71.90)	163 (82.3)	82.3
K-2 Words :	11 (10.2)	11 (9.09)	13 (6.6)	88.9
K-3 Words :	10 (9.3)	10 (8.26)	10 (5.1)	94.0
K-4 Words :	4 (3.7)	4 (3.31)	4 (2.0)	96.0
K-5 Words :	1 (0.9)	1 (0.83)	1 (0.5)	96.5
K-6 Words :	1 (0.9)	1 (0.83)	1 (0.5)	97.0
K-7 Words :				
K-8 Words :	1 (0.9)	1 (0.83)	1 (0.5)	97.5
K-9 Words :				
K-10 Words :				
K-11 Words :				
K-12 Words :				
K-13 Words :				
K-14 Words :				
K-15 Words :	1 (0.9)	1 (0.83)	1 (0.5)	98.0
K-16 Words :				
K-17 Words :				
K-18 Words :				
K-19 Words :				
K-20 Words :				
K-21 Words :				
K-22 Words :				
K-23 Words :				
K-24 Words :				
K-25 Words :				
Off-List:		??	4 (3.31)	4 (2.02)
Total (unrounded)	108+?	121 (100)	198 (100)	≈100.00

RELATED RATIOS & INDICES	
Pertaining to whole text	
Words in text (tokens):	198
Different words (types):	121
Type-token ratio:	0.61
Tokens per type:	01.64
Lexical density	0.55
Pertaining to onlist only	
Tokens:	194
Types:	117
Families:	108
Tokens per Family :	01.80
Types per Family :	01.08

4.3.2 Average scoring student's vocabulary profile, Group B

The average-scoring participant in Group B submitted an essay where word families from frequency levels K1 through K3 were represented. Moreover, one word from K5 and one word from K18 level were used. As such, the frequency levels represented in the essay do not differ significantly from that of the previous student. The first frequency level was represented even more numerous, with 166 tokens (81.8%). That number decreased to 18 (8.9%) in level K2 and 9 (4.4%) in K3. The next thousand words are represented by 3 tokens (K5 level) and 1 type (K18 level). The lexical density was 0.52, whereas the type-token ratio was 0.60. The productive vocabulary size of the participant was estimated to be 4,000 word families.

Table 17: The productive vocabulary text profile of the average scoring participant in research Group B.

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. Token (%)
K-1 Words :	82 (77.4)	91 (75.21)	166 (81.8)	81.8
K-2 Words :	14 (13.2)	15 (12.40)	18 (8.9)	90.7
K-3 Words :	7 (6.6)	7 (5.79)	9 (4.4)	95.1
K-4 Words :				
K-5 Words :	2 (1.9)	2 (1.65)	3 (1.5)	96.6
K-6 Words :				
K-7 Words :				
K-8 Words :				
K-9 Words :				
K-10 Words :				
K-11 Words :				
K-12 Words :				
K-13 Words :				
K-14 Words :				
K-15 Words :				
K-16 Words :				
K-17 Words :				
K-18 Words :	1 (0.9)	1 (0.83)	1 (0.5)	97.1
K-19 Words :				
K-20 Words :				
K-21 Words :				
K-22 Words :				
K-23 Words :				
K-24 Words :				
K-25 Words :				
Off-List:		??	4 (3.31)	6 (2.96)
Total (unrounded)	106+?	121 (100)	203 (100)	≈100.00

RELATED RATIOS & INDICES	
Pertaining to whole text	
Words in text (tokens):	203
Different words (types):	121
Type-token ratio:	0.60
Tokens per type:	01.68
Lexical density	0.52
Pertaining to onlist only	
Tokens:	197
Types:	117
Families:	106
Tokens per Family :	01.86
Types per Family :	01.10

4.4.3 Low scoring student's vocabulary profile, Group B

The lowest-scoring participant in Group B submitted an essay where only word families from frequency levels K1 through K4 were represented, with one word from level K11. The clear majority of tokens (189, or 92.2%) belonged to the first frequency level. The next levels (K2, K3 and K4) were represented by 6 (2.9%), 3 (1.5%) and 1 (0.5%) token respectively. The word from the last frequency level (*sup*, a contraction of *what's up?*) was meant as a slang expression, suggesting that it belonged to the wrong frequency level due to altered spelling. Moreover, the lexical density was 0.49, whereas the type-token ratio was 0.59. The productive vocabulary size of the participant was estimated to be 4,000 word families.

Table 18: The productive vocabulary text profile of the lowest scoring participant in research Group B.

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. Token (%)
K-1 Words :	93 (92.1)	105 (87.50)	189 (92.2)	92.2
K-2 Words :	4 (4.0)	4 (3.33)	6 (2.9)	95.1
K-3 Words :	2 (2.0)	3 (2.50)	3 (1.5)	96.6
K-4 Words :	1 (1.0)	1 (0.83)	1 (0.5)	97.1
K-5 Words :				
K-6 Words :				
K-7 Words :				
K-8 Words :				
K-9 Words :				
K-10 Words :				
K-11 Words :	1 (1.0)	1 (0.83)	1 (0.5)	97.6
K-12 Words :				
K-13 Words :				
K-14 Words :				
K-15 Words :				
K-16 Words :				
K-17 Words :				
K-18 Words :				
K-19 Words :				
K-20 Words :				
K-21 Words :				
K-22 Words :				
K-23 Words :				
K-24 Words :				
K-25 Words :				
Off-List:		??	5 (4.17)	5 (2.44)
Total (unrounded)	101+?	120 (100)	205 (100)	≈100.00

RELATED RATIOS & INDICES	
Pertaining to whole text	
Words in text (tokens):	205
Different words (types):	120
Type-token ratio:	0.59
Tokens per type:	01.71
Lexical density	0.49
Pertaining to onlist only	
Tokens:	200
Types:	115
Families:	101
Tokens per Family :	01.98
Types per Family :	01.14

4.5 Productive vocabulary profiles, Group C

In the following sections (4.5.1, 4.5.2 and 4.5.3), the high-, average- and low-scoring students' lexical profiles from group C are presented. The high-scoring student in this research group obtained a receptive vocabulary size of 10,600, the average-scoring student scored 8,400 word families, and the low-scoring student in the group scored 5,500 word families in the VST. In the following, their individual lexical profiles were presented.

4.5.1 High scoring student's vocabulary profile, Group C

The highest scoring student in research Group C had a receptive vocabulary size score of 10,600, with 5718 ms mean decision time. When it comes to the student's productive vocabulary size, words from the first six frequency levels were represented, except the K4 level. 166 tokens (83%) used in the essay belonged to the first frequency level, decreasing to 16 (8%) different tokens in the K2 level. The last three levels (K3, K5 and K6) were represented by 11, then only one word each (*dialect*, *infamous* and *denominator*). The type-token ratio was 0.57 and lexical density was 0.51. In general, 100 different word families are represented in the text. The productive vocabulary size of that participant was estimated to be 4,000 word families.

Table 19: The productive vocabulary text profile of the highest scoring participant in research Group C.

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. Token (%)
K-1 Words :	76 (76.0)	82 (71.93)	166 (83.0)	83.0
K-2 Words :	12 (12.0)	14 (12.28)	16 (8.0)	91.0
K-3 Words :	10 (10.0)	10 (8.77)	11 (5.5)	96.5
K-4 Words :				
K-5 Words :	1 (1.0)	1 (0.88)	1 (0.5)	97.0
K-6 Words :	1 (1.0)	1 (0.88)	1 (0.5)	97.5
K-7 Words :				
K-8 Words :				
K-9 Words :				
K-10 Words :				
K-11 Words :				
K-12 Words :				
K-13 Words :				
K-14 Words :				
K-15 Words :				
K-16 Words :				
K-17 Words :				
K-18 Words :				
K-19 Words :				
K-20 Words :				
K-21 Words :				
K-22 Words :				
K-23 Words :				
K-24 Words :				
K-25 Words :				
Off-List:		??	5 (4.39)	5 (2.50)
Total (unrounded)	100+?	114 (100)	200 (100)	≈100.00

RELATED RATIOS & INDICES	
Pertaining to whole text	
Words in text (tokens):	200
Different words (types):	114
Type-token ratio:	0.57
Tokens per type:	01.75
Lexical density	0.51
Pertaining to onlist only	
Tokens:	195
Types:	109
Families:	100
Tokens per Family :	01.95
Types per Family :	01.09

4.5.2 Average scoring student's vocabulary profile, Group C

The average-scoring participant in Group C had an estimated receptive vocabulary size of 8,200 word families, with a mean decision time of 10678 ms (nearly double the decision time of the highest-scoring participant). Frequency levels from K1 through K8 were represented. One word from level K5, K7 and K8 was used (*snob*, *pronouns* and *lingua*). 178 different tokens (89% of all tokens in the text), which belonged to the first frequency level were present in the text. There were 12 tokens used from the K2 level and 3 from K3. The type-token ratio was estimated at 0.51 and lexical density at 0.55. Words from 179 different word families were represented. The productive vocabulary size of the participant was estimated to be 4,000 word families.

Table 20: The productive vocabulary text profile of the average scoring participant in research Group C.

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. token
K-1 Words :	75 (83.3)	81 (80.20)	178 (89.0)	89.0
K-2 Words :	9 (10.0)	10 (9.90)	12 (6.0)	95.0
K-3 Words :	3 (3.3)	3 (2.97)	3 (1.5)	96.5
K-4 Words :				
K-5 Words :	1 (1.1)	1 (0.99)	1 (0.5)	97.0
K-6 Words :				
K-7 Words :	1 (1.1)	1 (0.99)	1 (0.5)	97.5
K-8 Words :	1 (1.1)	1 (0.99)	1 (0.5)	98.0
K-9 Words :				
K-10 Words :				
K-11 Words :				
K-12 Words :				
K-13 Words :				
K-14 Words :				
K-15 Words :				
K-16 Words :				
K-17 Words :				
K-18 Words :				
K-19 Words :				
K-20 Words :				
K-21 Words :				
K-22 Words :				
K-23 Words :				
K-24 Words :				
K-25 Words :				
Off-List:		??	3 (2.97)	4 (2.00)
Total (unrounded)	90+?	101 (100)	200 (100)	≈100.00

RELATED RATIOS & INDICES	
Pertaining to whole text	
Words in text (tokens):	200
Different words (types):	101
Type-token ratio:	0.51
Tokens per type:	01.98
Lexical density	0.55
Pertaining to onlist only	
Tokens:	196
Types:	98
Families:	90
Tokens per Family :	02.18
Types per Family :	01.09

4.5.3 Low scoring student's vocabulary profile, Group C

The lowest-scoring participant in the current group achieved a receptive vocabulary size of 5,500 word families, with a mean decision time of 3683 ms. This was also the shortest time spent by any participant in that group. Frequency levels from K1 through K3 were represented in the student's text, also with one word from each of the levels of K5 and K7 (*vicious* and *ghetto*). In total, 422 words were used in the text. The type-token ratio is 0,49 and lexical density was estimated at 0,50. Eighty five word families were represented in the student's writing, with 89,7% coming from the first frequency level. The productive vocabulary size of the participant was estimated to be 4,000 word families.

Table 21: The productive vocabulary text profile of the lowest scoring participant in research Group C.

RELATED RATIOS & INDICES	
Pertaining to whole text	
Words in text (tokens):	203
Different words (types):	99
Type-token ratio:	0.49
Tokens per type:	02.05
Lexical density	0.50
Pertaining to onlist only	
Tokens:	200
Types:	97
Families:	85
Tokens per Family :	02.35
Types per Family :	01.14

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. token (%)
				(%)
K-1 Words :	73 (85.9)	82 (82.83)	182 (89.7)	89.7
K-2 Words :	9 (10.6)	9 (9.09)	12 (5.9)	95.6
K-3 Words :	1 (1.2)	1 (1.01)	1 (0.5)	96.1
K-4 Words :				
K-5 Words :	1 (1.2)	1 (1.01)	1 (0.5)	96.6
K-6 Words :				
K-7 Words :	1 (1.2)	3 (3.03)	4 (2.0)	98.6
K-8 Words :				
K-9 Words :				
K-10 Words :				
K-11 Words :				
K-12 Words :				
K-13 Words :				
K-14 Words :				
K-15 Words :				
K-16 Words :				
K-17 Words :				
K-18 Words :				
K-19 Words :				
K-20 Words :				
K-21 Words :				
K-22 Words :				
K-23 Words :				
K-24 Words :				
K-25 Words :				
Off-List:		??	2 (2.02)	3 (1.48)
Total (unrounded)	85+?	99 (100)	203 (100)	≈100.00

4.6 Productive vocabulary profiles, Group D

In the following sections (4.6.1, 4.6.2 and 4.6.3), the high-, average- and low-scoring students' lexical profiles from group D are presented. The high-scoring student in this research group obtained a receptive vocabulary size of 9,700, the average-scoring student scored 7,600 word families, and the low-scoring student in the group scored 5,100 word families in the VST. In the following, their individual lexical profiles are presented.

4.6.1 High scoring student's vocabulary profile, Group D

The highest scoring participant in test group D obtained a receptive vocabulary size of 9,700 word families. The productive vocabulary size was 6,000 word families, with words from K1 through K5 represented in the text. One word from the K7 level was represented (*antagonist*). Since it is a literary term, one may assume that it was specifically targeted in the English classes and the student was able to recall it and use it correctly. 176 tokens from the first frequency level were used, covering 88% of the running words in the text. Nine tokens from the K2 level, 4 tokens from the K3 level, 1 from K4, and 7 from K5 level are used. The lexical density was 0.52 and the type-token ratio is 0.56.

Table 22: The productive vocabulary text profile of the highest scoring participant in research Group D.

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. Token (%)
K-1 Words :	84 (86.6)	95 (85.59)	176 (88.0)	88.0
K-2 Words :	5 (5.2)	6 (5.41)	9 (4.5)	92.5
K-3 Words :	3 (3.1)	3 (2.70)	4 (2.0)	94.5
K-4 Words :	1 (1.0)	1 (0.90)	1 (0.5)	95.0
K-5 Words :	3 (3.1)	3 (2.70)	7 (3.5)	98.5
K-6 Words :				
K-7 Words :	1 (1.0)	1 (0.90)	1 (0.5)	99.0
K-8 Words :				
K-9 Words :				
K-10 Words :				
K-11 Words :				
K-12 Words :				
K-13 Words :				
K-14 Words :				
K-15 Words :				
K-16 Words :				
K-17 Words :				
K-18 Words :				
K-19 Words :				
K-20 Words :				
K-21 Words :				
K-22 Words :				
K-23 Words :				
K-24 Words :				
K-25 Words :				
Off-List:		??	1 (0.90)	2 (1.00)
Total (unrounded)	97+?	111 (100)	200 (100)	≈100.00

RELATED RATIOS & INDICES	
Pertaining to whole text	
Words in text (tokens):	200
Different words (types):	111
Type-token ratio:	0.56
Tokens per type:	01.80
Lexical density	0.52
Pertaining to onlist only	
Tokens:	198
Types:	110
Families:	97
Tokens per Family :	02.04
Types per Family :	01.13

4.6.2 Average scoring student's vocabulary profile, Group D

The average-scoring participant in test group D obtained a receptive vocabulary size of 7,500 word families. The essay submitted by the participant included tokens from frequency levels K1 through K3. Three words from K5 level were represented (*emigrated, rags, nowadays*). With a gap at K4 level, the productive vocabulary size of this participant seems to be considerably smaller than that of the previous participant. 173 tokens from the first frequency level are used, covering 86.1% of the running words in the text. 11 tokens from K2 level, 11 tokens from K3 level, and 3 tokens from K5 level were used. The lexical density was 0.55 and the type-token ratio was 0.50. The productive vocabulary size of the participant was estimated to be 4,000 word families.

Table 23: The productive vocabulary text profile of the average scoring participant in research Group D.

RELATED RATIOS & INDICES	
Pertaining to whole text	
Words in text (tokens):	201
Different words (types):	101
Type-token ratio:	0.50
Tokens per type:	01.99
Lexical density	0.55
Pertaining to onlist only	
Tokens:	198
Types:	98
Families:	85
Tokens per Family :	02.33
Types per Family :	01.15

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. Token (%)
K-1 Words :	69 (81.2)	78 (77.23)	173 (86.1)	86.1
K-2 Words :	9 (10.6)	10 (9.90)	11 (5.5)	91.6
K-3 Words :	4 (4.7)	6 (5.94)	11 (5.5)	97.1
K-4 Words :				
K-5 Words :	3 (3.5)	3 (2.97)	3 (1.5)	98.6
K-6 Words :				
K-7 Words :				
K-8 Words :				
K-9 Words :				
K-10 Words :				
K-11 Words :				
K-12 Words :				
K-13 Words :				
K-14 Words :				
K-15 Words :				
K-16 Words :				
K-17 Words :				
K-18 Words :				
K-19 Words :				
K-20 Words :				
K-21 Words :				
K-22 Words :				
K-23 Words :				
K-24 Words :				
K-25 Words :				
Off-List:		??	3 (2.97)	3 (1.49)
Total (unrounded)	85+?	101 (100)	201 (100)	≈100.00

4.6.3 Low scoring student's vocabulary profile, Group D

The lowest-scoring participant in the current group achieved a receptive vocabulary size of 5,100 word families, with a mean decision time of 3683 ms. This was also the lowest receptive vocabulary size across all the research groups. One could expect this to be reflected in the productive vocabulary size, yet this was not the case. Frequency levels from K1 through K4 were represented in the student's text, with one word from level K6 (*contentment*). A relatively small coverage with words from K1 (79%) suggests use of words from the mid-frequency level. The type-token ratio was 0.49 and lexical density was estimated at 0.57. Eighty eight word

families were represented in the student's writing. The productive vocabulary size of the participant was estimated to be 5,000 word families.

Table 24: The productive vocabulary text profile of the lowest scoring participant in research Group D.

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. Token (%)
K-1 Words :	63 (71.6)	68 (69.39)	158 (79.0)	79.0
K-2 Words :	13 (14.8)	15 (15.31)	22 (11.0)	90.0
K-3 Words :	10 (11.4)	10 (10.20)	13 (6.5)	96.5
K-4 Words :	1 (1.1)	1 (1.02)	2 (1.0)	97.5
K-5 Words :				
K-6 Words :	1 (1.1)	1 (1.02)	1 (0.5)	98.0
K-7 Words :				
K-8 Words :				
K-9 Words :				
K-10 Words :				
K-11 Words :				
K-12 Words :				
K-13 Words :				
K-14 Words :				
K-15 Words :				
K-16 Words :				
K-17 Words :				
K-18 Words :				
K-19 Words :				
K-20 Words :				
K-21 Words :				
K-22 Words :				
K-23 Words :				
K-24 Words :				
K-25 Words :				
Off-List:		??	2 (2.04)	4 (2.00)
Total (unrounded)	88+?	98 (100)	200 (100)	≈100.00

RELATED RATIOS & INDICES	
Pertaining to whole text	
Words in text (tokens):	200
Different words (types):	98
Type-token ratio:	0.49
Tokens per type:	02.04
Lexical density	0.57
Pertaining to onlist only	
Tokens:	196

Types:	96
Families:	88
Tokens per Family :	02.23
Types per Family :	01.09

4.7 Lexical richness

After the essays were collected from the participating students and submitted into the Vocab Profiler Classic, the analysis started. Lexical richness is a general term that is considered when assessing effective vocabulary use. Lexical variation, lexical sophistication, lexical density, number of lexical errors, use of collocations, as well as the use of academic vocabulary, were analyzed (Read, 2000, p. 200). The program was used to gauge those measurements which may then be used to analyze the students' individual, overall progress in language acquisition, or in their lexical progress. In the present study, one of the objectives was to describe quantitatively the productive vocabulary size.

4.7.1 Lexical variation

TTR has proven to be sensitive to text length. The longer the text, the lower the TTR usually is (Gregori-Signes & Clavel-Arroitia 2015, p. 548). As mentioned in section 3.2, the data material had to be carefully considered before submitting them to the software. Essays were therefore shortened to 200 words and submitted into the Vocab Profiler to obtain the results. The TTR spanned from 0.49 (submitted by a low-scoring student) to 0.61 (submitted by a high-scoring student) in the participants' writing. A text with a TTR of 0.49 will have fewer types than one of 0.61. Generally, written texts have been found to have a TTR of 0.36-0.57 (Schmitt, 2000, p. 75). The submitted texts seem to be within that range at the bottom of the scale, but somewhat exceed the upper limit. It seems that there was a slight relationship between how well the student scored on the VST, and how high the TTR was obtained. Table 25 below shows the TTR for the high-, average- and low- scoring participants in the study:

Table 25: Type-token ratio of the high-, average- and low-scoring students in the study.

	TTR
Group A	
high	0.61
high	0.52
average	0.55
low	0.53
Group B	
high	0.61
average	0.60
low	0.59
Group C	
high	0.57
average	0.51
low	0.49
Group D	
high	0.52
average	0.55
low	0.49

4.7.2 Lexical sophistication

As mentioned in section 3.4.2, lexical sophistication is a subjective measure, and thus difficult to both generalize and compare the results (Laufer & Nation, 1995, p. 310). The instability of results is largely dependent on the fact that determining the number of advanced tokens in a text used for calculating lexical sophistication hinges on the researcher's own definition of advanced tokens. Such may be determined by the level of participants' education, their background, or the number of years or quality of exposure to the target language. However, defining too high requirements for the students may also affect the measurement. For instance, if 'advanced' words for the given corpus are defined to be from the K15 level of frequency, or a highly specialized vocabulary that the learners cannot possibly be familiar with, the LS ratio will be very low.

For the purpose of the present study, the advanced words have been defined as those belonging to frequency level K6 and above. The VP measures the coverage of the text with the various frequency levels, divided into thousands of word families. In addition, the coverage of words from the AWL is shown, together with off-list words (misspelled words, non-standard forms or proper nouns). These words do not need to be taken into consideration. An exception

is made if an academic word, included in the AWL, belongs to any of the frequency bands from K7 and above. The number of advanced tokens is then divided by the total number of lexical tokens. Another reason for choosing that specific threshold in defining advanced vocabulary is that the students' productive vocabulary size is on average approximately 7,000-word families. Word families above that band should be considered sophisticated for the given age group. The overview of LS of the targeted students is shown in the table below:

Table 26: Lexical sophistication of the high-, average- and low-scoring students.

Lexical sophistication	
Group A	
high	0.98 %
high	0.50 %
average	0.00 %
low	1.02 %
Group B	
high	1.03 %
average	0.00 %
low	0.50 %
Group C	
high	1.02 %
average	0.50 %
low	0.50 %
Group D	
high	0.51 %
average	0.00 %
low	0.00 %

The results spanned between 0% until 1.03%. with the group's average at 0.50%. Students with 0% LS had no words beyond the K6 frequency band. Students with approximately 1% had roughly one word per 100 tokens that was beyond the 6K frequency band in the running text. As mentioned above, this is a highly unstable measurement. For instance, if the threshold for advanced vocabulary would have been moved one frequency band up, to K8, the results would have differed greatly. Moreover, the ratio is affected by the number of repetitions. If a student uses very repetitive vocabulary, the number of tokens may be very low. If the student then uses a few words that are considered advanced. the LS ratio would be very high, even though the vocabulary of that student may hardly be considered generally sophisticated.

The results of the LS analysis would have been applicable in a comparative study, if a similar research study had been conducted. In a relatively homogenous student population. it may have some comparative value (Read, 2000, p. 204). Also, the same or a similar definition of lexical sophistication would make the results more stable.

4.7.3 Lexical density

The ratio spanned from 0.49 to 0.56 in the targeted students’ texts. There seems to be no relationship between the vocabulary size score in the vocabulary size test and the lexical density. The highest-scoring students did not always receive the highest score in lexical density. There was small variation between the density in the submitted texts. The following table shows the lexical density ratios for the high-, average- and low-scoring students in the targeted classes:

Table 26: Lexical density of the high-, average- and low-scoring students in the study.

	Lexical density
Group A	
high	0.5
high	0.54
average	0.5
low	0.56
Group B	
high	0.55
average	0.52
low	0.49
Group C	
high	0.51
average	0.55
low	0.51
Group D	
high	0.56
average	0.50
low	0.57

4.7.4 Lexical errors

The lexical errors categorized more closely in section 3.4.4. are summarized in the following table. The category of the mistake is given, and the mistake (in the sentence, with original

spelling) is then shown. The numbering in the left column refers to the categorization in section 3.4.4. The correct word choice or form is written in parenthesis:

Table 27: Types of lexical errors in the corpus.

Type of the mistake	The error in a sentence
II. 4 – Major spelling error	The <i>stander</i> family is... (standard).
I.A.2 – Incorrect. Semantically close	It's easy to <i>take</i> a hand on a gun... (put)
I.A.1 – Incorrect. Semantically unrelated	They are <i>though</i> from a young age... (taught)
I.B.1 - Combinations	A person <i>which</i> speaks... (who)
I.A.1 - Incorrect. Semantically unrelated	That <i>state</i> is a perfect example of... (statement)
I.B.2 - Phrases	Today's issue (...) <i>because that</i> people... (because of the fact that)
I.B.2 - Phrases	Many people want <i>to the old way</i> . (go back to the old ways)
I.A.2 – Incorrect. Semantically close	I believe americans all are very different. but they do have some common <i>abilities</i> . (qualities)
II.1 – Derivational errors	<i>Individuality</i> is thought of as a good ability to have. (Individualism)
I.B.1 – Two lexical items	They're also famous for their relationship with <i>everything food</i> . (everything that is edible. everything related to food)
II.1 – Derivational errors	Several factors are <i>including</i> . (included)
II.1 – Derivational errors	It includes progress and <i>succeed</i> . (success)
I.A.2 – Incorrect. Semantically close	Here he gets caught up in some trouble between two football <i>firms</i> . (clubs. gangs)
I.A.2 – Incorrect. Semantically close	Both these examples are about pay back and <i>getting</i> revenge. (seeking)

All of the types of lexical errors were represented in the corpus, with one exception: multiple errors with core lexical items were not found. Minor spelling mistakes (one misplaced or missing letter) were not taken into account. The students seemed to make more word choice errors than word form errors. Notably, none of the mistakes made by the students prohibited comprehension entirely. In all of the cases, it was unproblematic to find the actual word that was supposed to be used in the sentence.

4.7.5 Collocations

After performing the manual check, the following conclusions may be drawn. The most often used collocation type was adjective-noun collocation. Moreover, the number of collocations in each text was limited. Some of the collocations were used repeatedly by the students. The

limited number of collocations including prepositions (wrongfully used prepositions in collocations were not taken into account) may perhaps suggest that the students struggled with choosing the right preposition. In the table below, the collocations found in the students writing are presented.

Table 29: Types of lexical collocations used by the participants in Group A.

1	2	3	4	5	6	7	8	9	10
diversity of		speak English	many people		world language	hard-working		because of this	
		pursue idea	affluent society		lingua franca			all over the world	
			classless society		today's society				
			lower class						
			middle class						
			upper class						
			economic situation						
			ethnic background						
			social background						
			cultural background						
			British English						
			American English						
			inner neighborhoods						
			major city						
			young age						
			American dream						
			Indian English						

Table 30: Types of lexical collocations used by the participants in Group B.

1	2	3	4	5	6	7	8	9	10
people from	famous for		perfect example				move around		
hints of			Indian English				be surprised		
tons of			human races						
bunch of			American people						
exclusion of			American English						
relation to			British English						
			English speakers						
			varied language						

Table 31: Types of lexical collocations used by the participants in Group C.

1	2	3	4	5	6	7	8	9	10
		coin a term	upper class 3		mother tongue		care a lot		get out
		have a hard time	middle class 3		means of communication		work a lot		take care
			lower class 3				change drastically		
			enormous rate						
			first language						
			wide variety						
			vicious circle						
			a long time ago						

Table 32: Types of lexical collocations used by the participants in Group D.

1	2	3	4	5	6	7	8	9	10
Movie about		Achieve a dream	American people		Message of the movie		Get killed		To look after (oneself)
		Become rich	American dream 2						Get caught up (in trouble)
		Become famous	American identity						End up
		Seek revenge	American culture						
		Take place	Immigrant society						
		Want revenge	African-Americans						
		Belong somewhere	Different way						
			Hollywood industries						
			American society						
			Relevant theme						

4.7.6 Academic Vocabulary (AWL)

The table below shows the comprised summary of the targeted students' essays with regard to use of academic vocabulary, as presented in Coxhead's AWL. The table is divided into the research groups A, B, C and D, with the highest, average and low-scoring students' texts. The number of word families from the AWL, as well as types, are outlined. The coverage of the text oscillates between 1.76% (the lowest) to 4.38% text coverage with academic vocabulary.

Table 33: The academic vocabulary represented in the available corpora, as listed in AWL.

Corpus	Families	Types	Words from the AWL		
			Tokens	Text coverage in %	AWL families
A					
High	6	8	8	3.86%	Affect benefit corporate culture economy immigrate
High	1	1	1	0.48%	vary
Average	4	5	5	2.44%	achieve communicate normal pursue
Low	6	6	6	3.00%	culture diverse economy major philosophy vary
B					
High	8	8	9	4.57%	bulk dominant exclusion global minority transport varied versions
Average	7	7	9	4.37%	civil conflicts individuality issues text topics whereas
Low	2	2	2	0.98%	Issue tense
C					
High	7	7	7	3.50%	assume communicate constant diverse enormous global logic
Average	2	2	2	1.00%	migrate series
Low	1	1	1	0.49%	culture
D					
High	3	3	3	1.50%	relevant, seek, theme
Average	2	4	9	4.48%	achieve, immigrate
Low	12	14	18	9%	culture, diverse, ethnic, factor, identity, ideology, immigrate, individual, liberal, promote, region, reside

There does not seem to be a connection between the vocabulary size test score and usage of academic vocabulary. For instance, the lowest-scoring participant from group C only used seven different words from the AWL, making up for second to lowest text coverage (1.89%). Yet, so did the average-scoring participant from Group A, with seven academic words, and the lowest text coverage of 1.76%. The highest-scoring participants from Group A used 17 and 13 word families from AWL; from group B: 13 word families were represented from AWL; from group C, 22 word families were represented, and from group D 17 word families were represented from AWL. The average-scoring participants from group A used seven word

families, from group B: 19 word families. and from group C 12 word families. The lowest-scoring participants from group A used 11 word families, from group B: 15 word families. and from group C, 7 word families.

Some of the words were encountered repeatedly. ‘Communication’ or ‘communicate’. ‘immigrant’ or ‘immigrated’, ‘task’, ‘text’, ‘culture’ or ‘cultural’, ‘individual’ or ‘individuality’, ‘global’ were the words that were used most frequently. The reason may be that these words were specifically targeted during classes, that students were familiar and comfortable using them, or that the nature of the task affected the use of vocabulary.

4.8 The relationship between receptive and productive vocabulary knowledge

The distinction between the receptive and productive vocabulary size is generally accepted by both teachers, researchers and scholars (Read, 2000, p. 154). The assumption suggests that a learner of L2 first acquires receptive knowledge of vocabulary, and only later to acquire the skills to correctly use it, for instance in text production. Some of the aspects of the distinction between the two kinds of knowledge were discussed in section 2.6. This included the added difficulty when acquiring productive knowledge of a word, which then required stronger emphasis on developing the productive vocabulary size in language classroom, the fact that the receptive vocabulary size is bound to be larger than the productive one, as well as slower development of the productive vocabulary size in a foreign language, compared to that of receptive vocabulary size (as discussed in section 2.6 - Levitzky-Aviad & Laufer, 2013, p. 127).

The finding that the receptive vocabulary size is larger than the productive one, seems to be confirmed in the current study. Even though there were great differences between the participants, and across the groups, the receptive vocabulary size was found to be on average 8,338 word families. In comparison, the average productive vocabulary size in the sample was found to be 4,769 word families. However, the difference in productive size between the participants was not been as large (oscillating between 4.000-7.000 word families) as with the receptive vocabulary size. The students did not acquire the vocabulary at a great rate at this level. Neither did they seem to transfer words from their receptive to their productive vocabularies (Ozturk, 2015, p. 107).

The Vocab Profiler, Classic Edition has its shortcomings, since it does not account for potential collocational mistakes. Indeed, as mentioned in section 3.4.5, students would rather

use free lexical combinations rather than conventional collocations. Collocational knowledge is also a feature that distinguishes learners from native-like language production. From the list of the ten most commonly used collocations (summarized in Table 24 in section 3.4.5) the students in the sample size were most comfortable with using an adjective-noun type of collocation. The second most-frequently used collocation was the verb-noun type. There were many collocations that were hardly used in the sample size, such as adjective-preposition, noun-verb and preposition-determiner-noun type. Since developing conventionalized language forms is significant for the learners' communicative skills, these should be targeted specifically by the teachers (Levitzky-Aviad & Laufer, 2013, p. 129).

4.9 Brief summary of the results

A brief overview of the students' lexical profiles is presented in table above. The students with high-, average- and low scores on the receptive vocabulary size were chosen as a convenience extreme case sampling. In the third column to the left, the receptive vocabulary size of the given participant is given, followed by their estimated productive vocabulary size. The next four columns show the type-token ratio, lexical density, and what token coverage in percent of the running words in the text constitutes the high-frequency words (K1 frequency band, 0-1,000 word families).

There seems to be a pattern between the receptive and productive vocabulary size: the participants with a higher score on VST also showed larger productive vocabulary, with an exception in groups A and D, where the participant with the lowest score on VST actually had a higher productive vocabulary size than the average-scoring one. There was larger deviation in receptive vocabulary size across the sample, than in productive vocabulary size: the difference between highest and lowest scoring student in VST was 7,600 word families, while the difference between the highest and the lowest scoring student in VP was only 3,000 word families. This suggests that there may be considerable differences in comprehension across the student body. However, their productive skills did not differ considerably.

Additionally, the students seemed to use a relatively high percentage of high-frequency words. Indeed, the percentage of high-frequency words (K1 frequency band) oscillated between 79%-92.2%. The students seemed to rely on high-frequency vocabulary in their written production. Across groups A through C, the students with a high receptive vocabulary size use a smaller number of the high-frequency words. The low-scoring participant in group B had

92.2% of the running words in the text from that frequency level, meaning that 92.2% of the text was covered with K1 frequency level words (first 1,000 word families).

Table 34: Brief summary of the results of the individual lexical profiles of the participants.

	Participant	Receptive vocabulary size	Productive vocabulary size	TTR	LD	word families	K1 token coverage
Group							(%)
A	high	12.700	5.000	0.64	0.51	118	87.4
	high	12.700	4.000	0.61	0.53	110	84.6
	average	8.200	4.000	0.55	0.5	95	89.1
	low	6.400	5.000	0.53	0.56	95	83.5
B	high	12.700	7.000	0.61	0.55	108	82.3
	average	9.000	5.000	0.6	0.52	106	81.8
	low	6.200	5.000	0.59	0.49	101	92.2
C	high	10.600	4.000	0.57	0.51	100	83
	average	8.200	4.000	0.51	0.55	90	89
	low	5.500	4.000	0.49	0.5	85	89.7
D	high	9.700	6.000	0.56	0.52	97	88
	average	7.500	4.000	0.5	0.55	85	86.1
	low	5.100	5.000	0.49	0.57	88	79
AVERAGE		8.808	4.769	0.55	0.52	98.3	85.8

A slightly different picture is painted in the last group, where the participant with the lowest receptive vocabulary size score actually used the lowest percentage of high-frequency words (79%). This is usually a good sign. The low receptive vocabulary score could confirm this: not being able to match words with their meaning displays small vocabulary register. On the other hand, the student may have had low motivation to finish the test, which would also result in a low score. A relatively small number of word families in the text (88) may weigh against the former hypothesis.

Researchers report that the most frequent 2.000 words in English account for 80% of the running words in a text (Schmitt. Cobb et. al., 2015, p. 219), alternatively that almost 80% of the running words in a text are high-frequency words (Nation, 2013, p. 11). The participants' lexical profiles showed that the percentage of words from the first level of frequency was higher than that, even up to 92.2%. This may suggest that Vg1 students in Norway generally overuse the spoken language, even if presented with a formal task. Here, the register knowledge (as discussed in section 2.10) becomes relevant, since it is important to know which vocabulary the learners may not be aware of is appropriate for the situation.

5. Discussion of the results

5.1 Introduction

The following is a discussion of the results of receptive and productive vocabulary size research of teenagers in an upper secondary school in Norway. Then, the discussion of the productive vocabulary size and the main findings is presented. The findings are compared in the light of the research presented in Chapter 2, both in the Norwegian perspective and internationally. In the next section, implications and recommendations for teaching are described. Vocabulary learning and instruction are an important element of language learning. Methods and approaches that help facilitate better understanding and more optimal planning and implementation of the classes regarding vocabulary size findings are an important part of the research. Next, suggestions for further research are suggested. Finally, the limitations of the study are addressed.

5.2 Receptive vocabulary size

This study has set out to examine both receptive and productive English vocabulary size of Vg1 students in an upper secondary school in Norway, their characteristics, and the relationship between the two. The average receptive vocabulary size of students at Vg1 level in upper secondary school in Norway was found to be 8,338 words. Compared with Nation's (2006) overview of needed vocabulary size with 98% coverage of the text (see section 2.15), the average learner in this study can follow a daily conversation, watch a children's movie, read graded readers, and read novels such as *The Turn of the Screw*. Other activities, such as reading a newspaper or reading a more advanced novel, such as *Lady Chatterley's Lover*, would pose difficulty for some of the participants. However, this is only the case if 98% coverage of the text is considered ideal.

If the coverage threshold is lowered by only 1.3% to 96.7%, the vocabulary needed for the various activities is considerably lower (4,000 word families for a children's movie, 2,000 for a graded reader – Nation, 2006, p. 73-75). Less recognition of recurring words in a text may decrease comprehension and overall experience. Yet, the acquisition may still be enhanced if the learner is motivated and equipped with learning strategies. The recurring words, if targeted before they are encountered in the text, will seem familiar to the reader/listener (Nation, 2006,

p. 78). This requires attention from both the teacher to select the target glossary, and attention from the student's perspective. However, this is where vocabulary research and knowledge may facilitate language learning.

The average receptive vocabulary size suggests that the learners know the 2,000 most high-frequent words in the English language. These words comprised in the GSL (as discussed in section 2.15 – Read, 2000, p. 160), are the words that occur most often in English spoken and written texts. The current study shows that the learners are familiar with these words, meaning that they will be able to recognize and understand them. This is indeed important, as learners need to gain knowledge of the high-frequency vocabulary in order to be able to benefit from other aspects of language learning (Webb & Nation, 2017, p. 48). Moreover, the students need to further develop fluency, meaning the ability to process the words quickly, in the vocabulary they have already learned (Webb & Nation, 2017, p. 136). Engaging the fourth strand of language teaching (as discussed in section 2.11), the fluency development strand, is therefore crucial.

The large variation between the learners in the study should be pointed out. Among the whole population, the highest score was 12,700 word families (three students). This suggests much higher recognition of words than the average learner (8,338 word families) or the lowest-scoring students (5,100 word families). Such a large variation in receptive vocabulary knowledge should be accounted for when planning activities to conduct in the classroom. The more skilled learners may select unabridged books, newspapers, and other texts according to their level (Webb & Nation, 2017, p. 145-146). Ideally, the learners with lower scores should be provided with graded readers, 'easy-reading' texts and other adjusted material (Webb & Nation, 2017, p. 144). Moreover, the vocabulary targeted by the students should be different. The lower-scoring ones should be instructed on the high-frequency vocabulary to enhance their comprehension (Webb & Nation, 2017, p. 147). Otherwise, it is difficult to engage in activities involving meaning-focused input and output (Webb & Nation, 2017, p. 144). The highest-scoring participants should focus on less frequent words to improve their lexical accuracy, and more specialized and academic vocabulary (Webb & Nation, 2017, p. 145-146).

Compared to the students of other countries, the teenagers in Norway seem much more proficient when it comes to receptive vocabulary size. As discussed in section 2.16, the various groups of learners across the world have a considerably lower vocabulary size: 4,000 word families for Chinese English majors, 2,000 word families for Japanese and Omani EFL learners at University level, only 1,000 word families for French high school students and 1,200 word

families. for German high school students (Laufer, 2000, p. 48). In comparison, as mentioned in section 2.16, the EFL undergraduate students had a receptive vocabulary size score of 5,000 word families, whereas PhD students obtained 9,000 word families (Schmitt, 2010, p. 199). The very high score of Norwegian students may be attributed to the massive exposure to English outside the classroom (as discussed in section 2.4), as well as the long tradition of the communicative approach to language teaching in Norway (as discussed in section 2.5). It may be assumed that the comprehension level among the Norwegian learners is high.

5.3 Productive vocabulary size

The productive vocabulary size of the vgl students in the investigated classes in the upper secondary school was estimated to be on average 4,769 word families. Small differences across the sample were recorded, with the vocabulary size oscillating between 4,000 and 7,000 word families. Words used productively were analyzed using the Vocab Profiler, Classic Edition. The productive vocabulary size was shown to be significantly smaller than the receptive vocabulary size, which confirms the available research on this issue (Read, 2000, p. 154; Webb & Nation, 2017, p. 33; Mondria & Wiersma, 2004, p. 79; Levitzky-Aviad & Laufer, 2013, p. 127; Laufer & Goldstein, 2004, p. 427; Henriksen & Danelund, forthcoming). This is explained by the fact that recalling words receptively is substantially easier than actively recalling and producing them. Another important aspect is that productive learning involves receptive knowledge, whereas receptive learning does not involve aspects of knowledge needed for productive use, such as speaking and writing (Webb & Nation, 2017, p. 34).

However, the small differences in productive size across the sample show that the lexical development in productive use may be fairly similar among the students in the study. The students express themselves lexically in a similar manner. Indeed, especially research group D shows many similar features regarding the vocabulary size, LD, TTR, number of word families used and coverage of the text with K1 (high frequency) word families.

Such rigid distinction between the receptive and the productive vocabulary knowledge is not unproblematic, as research reports that it would be much more meaningful to consider the knowledge of a word on a continuum from receptive knowledge to fully operational productive knowledge of that word (Read, 2000, p. 154). Knowing a word in a productive way involves many aspects, such as spelling, meaning, pronunciation and

collocations. The distinction between a word's meaning, form and use are the aspects the learner's attention may be drawn to (Nation, 2005, p. 1). In the current research, meaning and use are focused on. If a student only knows the spelling and the meaning, but does not know in what collocations the word occurs, perhaps that student is not entirely familiar with the word. For instance, the student may know the word *monarchy*, know the meaning and spelling of the word. When describing a possibility of *monarchy* being *abolished*, the student may use the expression *stop monarchy*. This would be an evidence of not yet fully internalized collocations of the given word.

Sætevik defined sophisticated vocabulary as all low-frequency words from K10-K25 frequency level (Sætevik, 2018, p. 82) and reports an average LS ratio at 8th grade level in Norwegian lower secondary school to be 0.28%, whereas in 10th grade in the same study, the average LS was found to be 0.21%. This confirms the reported 'advanced' vocabulary in written texts, which is less than 1% among L2 learners (Nation, 2006, p. 79)

Another interesting aspect was the large coverage of the corpus with high-frequency words, coming from the K1 frequency band. As summarized in Ttable 30, the coverage of the running words in the text with K1 words could reach as much as 92.2%. The average coverage across the sample was 85.6%. This could suggest that the students in the current study use 'spoken', everyday language when expressing themselves in a more formal situation (school essay). Alternatively, they may lack the register knowledge. necessary for distinguishing between the different styles and registers of language. appropriate to the situation (Schmitt, 2000, p. 33). The various kinds of language proficiency, involving grammatical, textual, pragmatic and sociolinguistic knowledge, were more closely described in section 2.10).

An alternative explanation for the exceptionally high coverage of high-frequency words in the sample may be that the students are indeed familiar with more advanced vocabulary and less frequent words. Yet they do not feel comfortable using that vocabulary, are uncertain about the possible uses, collocations or meaning of the word, or they do not have the knowledge to recall the words productively, thus avoiding using them. This confirms the finding described in section 4.6 about the receptive and productive vocabulary being parts of a continuum on vocabulary knowledge, where the student gradually moves from knowing the word receptively, to developing a productive knowledge of the word (Read, 2000, p. 154).

Possibly, the students do not know mid-, low- frequency or academic vocabulary well enough to be able to use it. This is suggested by the comparison of the sample with AWL, as a

very small number of words were used, suggesting limited familiarity with this kind of vocabulary. This finding may imply that there is not nearly enough focus on the academic vocabulary in the second language classes the students were enrolled in, which indeed was confirmed in Skjelde's (2015) study discussed in section 2.14. According to Skjelde's findings, Norwegian textbooks used in English courses in Vg1 included only a limited number of academic vocabulary (Skjelde, 2015, p. 99-102). If the students are expected to master the academic vocabulary, they should be exposed to it through varied and systematic input and meaning-focused use.

Danelund and Henriksen (2015) report similar findings regarding the receptive and productive vocabulary size of students in the first level at an upper secondary school in Denmark. These included relying on the K1 and K2 frequency level words in written production, as well as exhibiting limited LS and LD in addition to limited and erroneous usage of collocations (Henriksen & Danelund, 2015, p. 36; Henriksen, 2013, p. 35-36). Several factors may account for these features. Learners may use avoidance strategies, relying solely on the high-frequency words which they have mastered well (Henriksen & Danelund, 2015, p. 51). Despite continuous vocabulary acquisition, development of lexical sophistication in free writing stagnates, leading to fossilization. Lack of incentive to use more sophisticated vocabulary may explain that development (Henriksen & Danelund, 2015, p. 52).

Having reached the plateau of productive vocabulary, the students may consider themselves proficient and able to communicate according to their needs. Moreover, the teachers may not have pushed the students to use more sophisticated vocabulary in their output (Henriksen & Danelund, 2015, p. 50). This may be one of the reasons why the students in the current study showcase much larger receptive vocabulary, than productive vocabulary. The 'play-it-safe' strategy may result in faster writing or smaller number of errors (Henriksen & Danelund, 2015, p. 29). The high-frequency words constitute thus the large chunks of the students' written production, even though the research shows that the students have the potential for understanding and use of more precise, low-frequency vocabulary.

When it comes to use of collocations, Durrant and Schmitt (2008) report that adult learners of English as L2 indeed experienced retaining collocations from large exposure of meaning-focused input (Durrant & Schmitt, 2008, p. 182). The lack of knowledge of formulaic language may be attributed to a lack of instruction, but rather to missing input in the language instruction (Durrant & Schmitt, 2008, p. 163). Yet, the most significant improvement in

collocation-use was seen after fluency-oriented repetition, rather than after exposing the students to same collocations, only in various contexts (Durrant & Schmitt, 2008, p. 182).

5.4 Teaching implications and recommendations

An important feature of the English subject curriculum in Norway is that it does not specify methods for teaching English (Skulstad, 2018, p. 57). However, there is an emphasized interest in communicative competence. Newly-gained understanding of the students' vocabulary size should bear implications for language teachers. Teaching should involve a systematic development of vocabulary knowledge with attention to spelling, collocations and core meanings. This will eventually lead to independent vocabulary use, but only if both productive and receptive use are promoted through fluency development (Webb & Nation, 2017, p. 37). Another implication includes use of context. As discussed in section 5.1, students use many types of collocations to a minimal extent, which may impair their communicative competence. Oral and listening skills are enhanced by collocational knowledge, in addition to reading speed (Gitsaki, 1999, p. 28). Moreover, problems of style and usage may be decreased by teaching collocations. Providing students with context is one of the best ways of reducing the learning burden and enabling one to later recycle the familiar vocabulary in proper way.

Another implication for the teachers involves mid- and low-frequency words. Having knowledge about the students' excessive dependence on high-frequency words in text production, the teachers should create opportunities for exposure to mid- and low-frequency vocabulary, as well as academic vocabulary, in a contextualized environment. Meaningful input that includes target vocabulary may involve journals and magazines written for native speakers and documentaries and other factual materials that target vocabulary above K3 frequency level. As a means of engaging the second and fourth strand of language teaching, which are meaning focused output and fluency development (discussed in section 2.9), the students should also be provided with opportunities to use that vocabulary productively in meaningful, both spoken and written activities.

The awareness of the need for generative use of taught vocabulary is crucial in language instruction (Nation, 2013, p. 110). The students do seem to use limited vocabulary in writing, which may suggest that they need guidance on productive use of vocabulary. Generative processing, either receptive or productive, may help in acquiring new vocabulary. The

vocabulary has to be processed in new contexts and used in other ways, in order for the student to learn it. Again, it is the teacher's responsibility to provide meaningful activities and opportunities for the student to generatively process the word, until it is retained and may be recalled and produced correctly.

The small difference in productive vocabulary size, paired with the significant difference in receptive vocabulary size across the sample bring other implications. It may mean that the students with the lowest receptive vocabulary sizes (5,000-7,000 word families) struggle with understanding the class material, or other kinds of communication in target language. The importance of differentiated input should thus be highlighted (Webb & Nation, 2017, p. 145). Graded readers, shortened and simplified material that is accessible for students should be a part of the teaching material. Nation (2014) argues for the use of a diverse corpus, where texts of different genres and within different topics should be used by the learners in order to meet the most words (Nation, 2014, p. 12). This applies to students of all needs and proficiency levels.

Another aspect which may have implications for teachers was the limited lexical richness, measured by lexical sophistication, density and variation in the students' free writing. There seems to be little emphasis and incentives in teaching for students to explore more precise and sophisticated lexical structures (Henriksen & Danelund, 2015, p. 49). Indeed, greater growth of vocabulary may be ensured by larger exposure to meaning-focused input and output through extensive reading and writing. Yet, stronger implementation of requirements in terms of assessment should be required (Ozturk, 2015, p. 107). Extra support in learning and taking in use more advanced vocabulary should be required. Adjusting the assessment criteria, from focusing on the communicative aspects of the text, to include more lexical precision, would focus students' attention on using more advanced vocabulary. If the teacher's attention is especially focused towards the communicative skills, the lexical precision may suffer. Including systematic focus and feedback on the lexical aspect of the students' written production may help alleviate this issue, as well as give the students motivation to start using vocabulary from mid- and low-frequency levels, as well as other linguistic structures (Henriksen & Danelund, 2015, p. 51). Moreover, using diagnostic language tests should be assigned larger importance, as it may help teachers in their instruction and understanding of the students' needs and current lexical proficiency. There is currently no tradition for applying diagnostic tests for L2 learners in upper secondary schools in the Scandinavian perspective (Henriksen & Danelund, 2015, p. 51).

Deliberate focus on vocabulary teaching should have specific areas of interest, such as collocations, high-, mid- and low-frequency vocabulary, in order to help students increase the usage of more precise vocabulary, mid- and low-frequency words as well as academic vocabulary. For instance, the use of collocations explored in this study shows that the students use particular types of collocations to a limited extent, often ignoring formulaic language and using free combinations instead. It is argued that collocations should not be left for incidental acquisition but offered deliberate attention in class (Webb & Nation, 2017, p. 75). Since most collocations are low-frequency in occurrence, vast amounts of input would be required for the students to encounter them (Webb & Nation, 2017, p. 75).

Indeed, deliberate focus on vocabulary learning should be part of a well-balanced vocabulary instruction. However, the attention to learning through differentiated input should not be forgotten (Nation, 2014, p. 13). Being exposed to large amounts of meaning-focused input in the L2 is crucial for lexical development. Resources that may help facilitate incidental vocabulary learning are many, including: written input at the appropriate level that answers to the students' needs and interests, spoken input, television, movies or TED talks (Webb & Nation, 2017, p. 218-220). The resources should be found, designed and updated accordingly in relation to technological advances and time. In an ideal institution, teachers can cooperate to create a resource base of useful vocabulary for the students.

5.5 Limitations of the study

The main limitation of the study is the low number of participants. Since the number of participants is limited, the results may not be generalized, nor can any overall conclusions that may apply to the whole population be. The representativeness of the study is thus compromised to a large degree. This is also due to the fact that the participants in the current study are enrolled in a general studies program. Suggestions for minimising some of these drawbacks are presented in the next section: Suggestions for further research. A larger sample size may have given different results, alternatively more reliable results concerning the students' receptive and productive vocabulary size, as well as the lexical features of their writing and use of academic vocabulary.

Furthermore, the study was not designed to measure the vocabulary size of English as third language, nor what influence does the multilingual background have on vocabulary

growth in English. It would have been ideal to include how the mother tongue influences learning of English, whether the additive nature of language may be accounted for in the acquisition of vocabulary. In order to carry out such an analysis, a more complex tool would need to have been applied. There is a reason to expect that multilingual background influences learning and understanding in the target language as well, and this should thus be accounted for.

Other limitations may be inherent in the statistical measures used during the research. As discussed in Chapter 3, both TTR, which accounts for lexical variation, and LS, are flawed to some degree. Using more reliable measurements, with more objective characteristics, could help alleviate this issue. For instance, Covington (2008) suggests using moving-average TTR instead of a regular one, to limit its sensitivity to length (Covington, 2008, p. 1). Another aspect is basing the research on the widely used notion of word frequency: words are more than just how often they are encountered in use (Covington, 2008, p. 2). Even more reliable statistical measurements are necessary to establish a valid picture of learners' lexical proficiency. It is also the reason why, in the current study, qualitative measurements were employed, alongside the quantitative measurements.

5.6 Suggestions for further research

Possible research perspectives within the field of vocabulary acquisition and vocabulary assessment are many. Language teaching benefits greatly from applying research-based methods and approaches. The analysis of written language may facilitate teachers' perception of language proficiency, which may then be used for diagnosis, planning and assessment. Foreign language instruction in general has developed significantly in the past century, as discussed in section 2.3, mostly due to changing perspectives on teaching and needs of the learners. Knowledge about vocabulary specifically has only recently been given more attention, as it has become recognized as a vital part of language teaching and learning. Thus, the current study aimed at studying the receptive and productive vocabulary size of upper secondary school students in Norway. Sætevik's findings (2018) have been an important contribution in the field, where lower secondary students in 8th and 10th grade were targeted. Replication of both studies could either validate the results or discard them. Studies of the same field, only in other parts of the country could show whether the studies were representative for the country and the whole population.

Moreover, studies on vocabulary teaching methods should be a priority. Finding the methods that facilitate learning in the best way and ensure the most optimal retention is crucial. Also, research concerning strategies that may be applied when working with a highly heterogeneous and/or homogeneous student mass should be a priority. Teaching a large class (the classes in which the current study was conducted counted up to 30 students each) may pose challenges in terms of providing individualized feedback, they may reduce teacher-student interaction, difficulties in monitoring comprehension and progress or aversion to partake in class in front of many students (Webb & Nation, 2017, p. 141). There are several ways to diminish these challenges, for instance by providing plenty of opportunities for group work. Activities involving meaning-focused input or promoting student engagement may be crucial for dealing with a large student mass (Webb & Nation, 2017, p. 142-143). Finding the right ways to facilitate vocabulary learning should be research-based and adjusted to the size, proficiency, motivation and needs of the students.

An important link should be made between instruction and writing development. It has been established that receptive vocabulary size is larger than productive vocabulary size. Research should be conducted regarding the kind of instruction that may help students operationalize vocabulary they know receptively. To be able to use productively, should be applied. Also, teachers should be equipped with strategies which facilitate assessing that development. Moreover, instructional approaches that support lexical proficiency improvement should be emphasized and thoroughly researched.

As Skjelde's (2015) study suggests, the textbooks used in the first grade of upper secondary school in Norway make use of academic vocabulary to a limited extent. Having investigated multiple textbooks commonly used in Norway, the conclusion was unanimous. Yet, the students are required to know, learn and produce texts with academic vocabulary and formal register. It is therefore vital that educators provide them with meaningful input consisting of such vocabulary on a regular basis. This may indeed change when the new national curriculum takes effect in autumn 2021 (Utdanningsdirektoratet, 2018a). New textbooks will need to be written and published. Hopefully, the input will involve the vocabulary that the students are required to use in the higher levels of education. Research could be devised to investigate what kind of vocabulary is covered, from which frequency levels or from which specialized field, and what kind of vocabulary tasks are there.

As mentioned in section 2.16, the Danish studies (Henriksen & Danelund, 2015; Henriksen, 2013) report that students at the first year of upper secondary school have achieved

receptive and productive vocabulary sizes that are hardly satisfying. As reasons for such development, several factors are reported, among others avoidance strategies, small motivation for expanding the vocabulary, and reliance on high-frequency vocabulary (Henriksen & Danelund, 2015, p. 50). Teachers' attention to deliberate vocabulary instruction was limited, since the communicative aspect of the students' production and the fluency of communication was given more attention (Henriksen & Danelund, 2015, p. 49). Research on teachers' attitudes toward deliberate and incidental vocabulary teaching and their influence on students' approaches towards expanding their vocabulary would be of interest. Alternatively, research may focus on how much time in an English language classroom is devoted to vocabulary instruction, vocabulary learning strategies (such as using dictionary or inferencing) and vocabulary expansion tasks. Also, the learners' attitudes and perceptions regarding own vocabulary acquisition should be focused on.

Lastly, research is necessary to study the vocabulary knowledge of vocational students. The current research targeted general studies classes. Even with the pronounced differences in future uses, both programs use the same national curriculum and are supposed to pass similar exams. After the new curriculum is introduced, this will change, and the vocational students will have more profession-related learning with specialized vocabulary in their field. It will be important to see what approaches teachers in vocational programs adopt teaching English, what methods are applied, how large is the vocabulary size of the students is and what vocabulary is targeted in the classroom.

6. Conclusion

The current study has investigated the receptive and productive vocabulary size of in English of Vg1 students (first grade of upper secondary school, or year 11) in Norway. The target group consisted of 100 learners from academic studies in upper secondary school in Rogaland, Norway. The research was based on data material gathered using an online vocabulary size test available at my.vocabularysize.com devised at Victoria University, Wellington, which measured the receptive vocabulary size as well as Vocab Profiler (VP), a data-driven software available online, which measured the productive vocabulary size. The students' essays were submitted to VP for quantitative analysis of their students' written production.

The receptive vocabulary size was found to be on average 8,338 word families. This means that the students had on average a receptive vocabulary size large enough to achieve 98% text coverage in order to conduct all the activities described in Table 3 in section 2.16 (Nation, 2006, p. 79). This entails reading a novel, reading a newspaper, reading graded readers, watching a children's movie, and following unscripted spoken English. The highest score was 12,700 word families, whereas the lowest score was 5,100 word families.

The receptive vocabulary size of the Norwegian teenagers enrolled in the first grade of upper secondary school was shown to be substantially larger than that of EFL students in other countries, both at high school and university level, as summarized in Table 4 in section 2.16. This may be due to the large emphasis on communicative skills in English teaching in the Norwegian school, which promotes meaningful input in everyday situations through the use of the target language. In addition, the massive exposure to English in extramural activities, as well as the high, almost second language status of English (as discussed in section 2.4) in Norway, may be some of the explanations.

The average productive vocabulary size was estimated to be 4,769 word families. This confirms that the receptive vocabulary size was larger than the productive vocabulary size. The differences in the vocabulary size across the sample were not as large as in the receptive vocabulary size. There was large coverage of the first 1,000 word families, which suggests little familiarity or little productive knowledge of less frequent or academic words. The participants used some collocations, especially the adjective-noun type. Some of the ten most used types of collocations are not represented in the corpus. This may indicate that students used free word

combinations rather than fixed expressions. This feature impairs their communicative skills and distinguished them from native-speakers.

The lexical richness of the students' writing was measured by lexical variation, density and sophistication. In addition, the written production has been analyzed regarding use of collocations, academic vocabulary and lexical errors. The average type-token ratio, which is a measure of lexical variation, was estimated to be 0.55. Average lexical density was found to be 0.52. On average, 98.3 word families were found in the individual texts. 85.8% of the texts were covered with K1 frequency words, meaning high reliance on the high-frequency. Low lexical sophistication, defined from K6 and above, was found. The students used, to a large extent, high-frequency words, which may be explained with 'playing-it-safe' strategy described above. Academic vocabulary as presented in the AWL was used sparsely.

7. References

- Bryman, A. (2001). *Social Research Methods*. Oxford: Oxford University Press.
- Brevik, L.M. & Hellekjær, G.O. (2017). Outliers: Upper Secondary School Students Who Read Better in the L2 than in L1. *International Journal of Educational Research* (2017). <http://dx.doi.org/10.1016/j.ijer.2017.10.001>.
- Coady, J. & Huckin, T. (1997). *Second Language Vocabulary Acquisition*. Cambridge: University Printing House.
- Cobb, T. (2002). *Web Vocabprofilers*. Retrieved from <http://www.VP.ca/vp/>. an adaptation of Heatley, Nation & Coxhead's (2002) *Range*.
- Council of Europe. (2001). *Common European Framework of Reference: Learning, Teaching, Assessment*. Retrieved from: <https://rm.coe.int/1680459f97>.
- Covington, M. (2008). *Text Statistics*. Retrieved from: <http://www.covingtoninnovations.com/mc/8570/TextStatistics.pdf>
- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34(2): 213-238.
- Davis, M. & Gardner, D. (2010). *Word Frequency List of American English*. Available online at www.wordfrequency.com
- Dornyei, Z. (2011). *Research Methods in Applied Linguistics*. Oxford: Oxford University Press.
- Durrant, P. & Schmitt, N. (2008). Adult Learners' Retention of Collocations from Exposure. *Second Language Research* 26,2 (2010); pp. 163-188).
- Ellis, R. & Shintani, N. (2014). *Exploring Language Pedagogy Through Second Language Acquisition Research*. Oxon: Routledge.
- Eyckmans, J. (2004). *Measuring Receptive Vocabulary Size Reliability and Validity of the Yes/No Vocabulary Test for French-speaking Learners of Dutch*. https://www.lotpublications.nl/Documents/91_fulltext.pdf (retrieved on 13.09.2018).
- Gray, D. (2009). *Doing Research in the Real World*. London: Sage Publications. 2nd ed.
- Gitsaki, C. (1999). *Second Language Lexical Acquisition. A Study of the Development of Collocational Knowledge*. Bethesda: International Scholars Publications.
- Golden, A. & Kulbrandstad, L.A. (2018). *A1. B1. B2 ... - <<den nye tellemåten>> på Språklæringsfeltet*. In Danbolt, A.M., Alstad, G. & G. Randen (2018). 'Litterasitet og flerspråklighet'. Bergen: Fagbokforlaget.
- Gregori-Signes, C. & Clavel-Arroitia, B. (2015). *Analysing Lexical Density and Lexical Diversity in University Students' Written Discourse*. *Procedia – Social and Behavioral Sciences* 198 (2015) 546-556.
- Hellekjær, G. O. (2019). Læreplan 2020: Engelsk Programfag – Noen Forslag til Endringer. *Bedre Skole* nr. 1. 2019 – 31. årgang.
- Hellekjær, G. O. (2008). A Case for Improved Reading Instruction for Academic English Reading Proficiency. *Acta Didactica*. vol. 2. nr. 1. art. 3, p. 1-17.
- Henriksen, B. & Danelund, L. (2015). "Studies of Danish L2 Learners' Vocabulary Knowledge and the Lexical Richness of their Written Production in English". In Pietilä, P., K. Doró, & R. Pipalová (Eds.), *Lexical issues in L2 writing*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Henriksen, B. (2013). 'Research on L2 Learners' Collocational Competence and Development - A Progress Report'. *L2 vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis*. 29-56.
- Hirsh, D. (2010). *Researching Vocabulary*. Continuum Companion to Research Methods in Applied Linguistics. 222-239.
- Jaen, M. M. (2017). *List of Frequent Collocations*. University of Granada. Retrieved from: http://www.ugr.es/~inped/documentos/lista_coloc_freq.pdf.

- Koizumi, R. (2012). Relationships between Text Length and Lexical Diversity Measures: Can We Use Short Texts of Less than 100 Tokens? *Vocabulary Learning and Instruction*. 1(1). 60-69. DOI: 10.7820/vli.v01.1.koizumi
- Laufer, B. & Nation, P. (1995). ‘Vocabulary Size and Use: Lexical Richness in L2 Written Production’. *Applied Linguistics*. Volume 16. Issue 3. 1 September 1995. Pages 307–322. <https://doi.org/10.1093/applin/16.3.307>.
- Laufer, B. & Nation, P. (1999). ‘A Vocabulary-Size Test of Controlled Productive Ability’. *Language Testing*. 16 (1) 33-51.
- Laufer, B. & Goldstein, Z. (2004). ‘Testing Vocabulary Knowledge: Size, Strength, and Computer Adaptiveness’. *Language Learning. A Journal of Research in Language Studies*. Volume 54. Issue 3. September 2004. Pages 399-436. <https://doi.org/10.1111/j.0023-8333.2004.00260.x>.
- Laufer, B. (2000). Task Effect on Instructed Vocabulary Learning: The Hypothesis of ‘Involvement’. *Selected papers from AILA '99 Tokyo* (pp. 47-62). Tokyo: Waseda University Press.
- Levitzky-Aviad, T. & Laufer, N. (2013). *Lexical Properties in the Writing of Foreign Language Learners over Eight Years of Study: Single Words and Collocations*. EUROSLA MONOGRAPHS SERIES 2 L2 vocabulary acquisition. knowledge and use. (pp. 127-148). Retrieved from: <http://www.eurosla.org/monographs/EM02/Levitzky-Aviad-Laufer.pdf>.
- Kumaravadivelu, B. (2005). *Understanding Language Teaching. From Method to Postmethod*. New Jersey: Routledge.
- Mack, N., C. Woodsong, Macqueen. K. M., Guest. G. & E. Namey. (2005). *Qualitative Research Methods: A Data Collector's Field Guide*. North Carolina: Family Health International.
- Mondria, J.A. & Wiersma, B. (2004). Receptive. Productive. and Receptive + Productive L2 Vocabulary & Learning: What Difference Does it Make? In: *Vocabulary in a second language: selection, acquisition, and testing*. Ed. by Paul Bogaards, Batia Laufer, p. 790-100. Amsterdam: John Benjamins Publishing Company.
- McCarthy, P.M. & Jarvis, S. (2010). Mtd. vocd-D. and HH-D: A Validation Study of Sophisticated Approaches to Lexical Diversity Assessment. *Behaviour Research Methods*. 42. 381-392. DOI: 10.3758/BRM.42.2.381
- McGarrighan, D. & Rugesæter, K. P.N. (2018). *Of Course You Can. An English Language Course for Teachers*. Bergen: Fagbokforlaget.
- Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury House.
- Nation, P. & Newton. J. (1997). *Teaching Vocabulary*. Cambridge: Cambridge University Press.
- Nation, P. (2014). How Much Input Do You Need to Learn the Most Frequent 9.000 Words? *Reading in a Foreign Language*, October 2014. Volume 26. No. 2 ISSN 1539-0578 pp. 1–16.
- Nation, I. S. P. & Anthony. L. (2013). Mid-frequency readers. *Journal of Extensive Reading*. 1. 5–16.
- Nation, I. S. P. (2013). *Learning Vocabulary in Another Language* (second ed.). Cambridge: University Printing House.
- Nation, P. & Yamamoto. A. (2012). Applying the Four Strands to Language Learning. *International Journal of Innovation in English Language Teaching*. Volume 1. Number 2. pp. 167-181.
- Nation, I.S.P. (2007). ‘My Ideal Vocabulary Teaching Course’. New Zealand: Victoria

- University of Wellington. Retrieved from:
<https://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/my-ideal-vocabulary-teaching-course.pdf>.
- Nation, I.S.P. & Beglar, D. (2007). A Vocabulary Size Test. *The Language Teacher* 31(7): 9-13.
- Nation, I.S.P. (2006). How Large a Vocabulary is Needed for Reading? *The Canadian Modern Language Review/La Revue Canadienne des langues vivantes*. 63.1 (September).
- Nation, P. (2005). ‘‘Teaching Vocabulary’’. *Asian EFL Journal*. Retrieved from:
http://www.asian-efl-journal.com/sept_05_pn.pdf
- Nguyen, L. and Nation, I.S.P. (2011). A bilingual vocabulary size test of English for Vietnamese learners. *RELC Journal*. 42(1). 86-99.
- Nation, I.S.P. & Chung, T.M. (2003). Technical Vocabulary in Specialized Texts. *Reading in a Foreign Language*, Volume 15, Number 2, October 2003
- Nizonkiza, D. (2016). First-Year University Students’ Receptive and Productive Use of Academic Vocabulary. *Stellenbosch Papers in Linguistics*. Vol. 45. 2016. 169-187
doi: 10.5774/45-0-215
- Nordby, A. (2007). *En Empirisk Studie av Sammenhengen Mellom Vokabular og en Rekke Kognitive og Lingvistiske Ferdigheter*. University of Oslo. Retrieved from:
<https://www.duo.uio.no/bitstream/handle/10852/30975/masteroppgave.anna.nordby.pdf?sequence=1>.
- Opdal, T.S. (2017). *Academic and General English in Norwegian Upper Secondary School. A Study of Students’ Proficiency and Ability to Distinguish between the Two Language Styles*. Norwegian University of Science and Technology Faculty of Humanities Department of Language and Literature. Retrieved from:
<https://brage.bibsys.no/xmlui/bitstream/handle/11250/2463516/Masteroppgave%20Thea%20Opdal.pdf?sequence=4>.
- Orosz, A. (2009). *The Growth of Young Learners’ English Vocabulary Size: Processes and Outcomes*. 10.21832/9781847691477-015.
- Ozturk, M. (2015). Vocabulary Growth of the Advanced EFL Learner. *The Language Learning Journal*, 43:1, 94-109.
- Paribakht, T. S. & Wesche, M. (eds.) (2003). *Vocabulary Enhancement Activities and Reading for Meaning in Second Language Vocabulary Acquisition*. Cambridge: Cambridge University Press.
- Richards, J.C. & Rodgers, T.S. (2015). *Approaches and Methods in Language Teaching*. Cambridge: Cambridge University Press.
- Read, J. (2000). *Assessing Vocabulary*. New York: Cambridge University Press.
- Schmitt, N., Cobb, T., Holst, M. & D. Schmitt. (2015). ‘‘Replication Studies. How Much Vocabulary is Needed to Use English? Replication of van Zeeland & Schmitt (2012).
- Schmitt, N. & Schmitt, D. (2014). A Reassessment of Frequency and Vocabulary Size in L2 Vocabulary Teaching. *Language Teaching*. Volume 47/Issue 04/October 2014. pp. 484-503.
- Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Basingstoke: Palgrave Macmillan.
- Schmitt, N. (2008). Review Article: Instructed Second Language Vocabulary Learning. *Language Teaching Research*. 2008: 12: 329.
- Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
- Schmitt, N. (1997). *Using a Word Knowledge Framework to Research Vocabulary*. Doctoral Thesis. University of Nottingham.

- Schmitt, N. (1995). A Fresh Approach to Vocabulary: Using a Word Knowledge Framework. *Relc Journal*. Vol. 26. No. 1. June 1995.
- Shin, D. & Nation, P. (2007). Beyond Single Words: the Most Frequent Collocations in Spoken English. *E LT Journal* Volume 62/4 October 2008.
- Skjelde, K. (2015). *Academic Vocabulary. Unleashed potential? A corpus study of English course materials for advanced Norwegian learners of English*. Master's thesis. University of Bergen.
- Skoglund, D.E. (2006). *A Comparison of Norwegian and American Students' English Vocabulary Usage in Upper Secondary Schools*. University of Oslo. Retrieved from: https://www.duo.uio.no/bitstream/handle/10852/32424/Dallas_Skoglund_Master_Thesis.pdf?sequence=2.
- Skulstad, A.S. (2018). Communicative Competence. In A.-B. Fenner & A.S. Skulstad. *Teaching English in the 21st Century*. Bergen: Fagbokforlaget.
- Sreehari, P. (2012). Communicative Language Teaching: Possibilities and Problems. *English Language Teaching*. Vol. 5. No. 12; 2012. pp. 87-93.
- Sætevik, J.A. (2018). *English Vocabulary Size in Lower Secondary School*. Master's thesis. University of Bergen.
- Utdanningsdirektoratet. (2019). *Proposal for the Renewal of English Subject Curriculum*. <https://hoering.udir.no/Hoering/v2/271?notaId=512>.
- Utdanningsdirektoratet. (2018). *Overall Part of the Curriculum*. Retrieved from: <https://www.udir.no/laring-og-trivsel/lareplanverket/overordnet-del/>.
- Utdanningsdirektoratet. (2013a). *Competence Aims after Year 10. English Subject Curriculum*. <https://www.udir.no/kl06/ENG1-03/Hele/Kompetansemaal/competence-aims-after-year-10?lplang=http://data.udir.no/kl06/eng>.
- Utdanningsdirektoratet. (2013b). *Purpose of English*. <https://www.udir.no/kl06/ENG1-03/Hele/Formaal?lplang=http://data.udir.no/kl06/eng>.
- Utdanningsdirektoratet. (2011). *Core Curriculum for Primary, Secondary and Adult Education in Norway*. https://www.udir.no/globalassets/filer/lareplan/generell-del/core_curriculum_english.pdf.
- Webb, S. & Nation, P. (2017). *How Vocabulary is Learned*. Oxford: Oxford University Press.
- Zhang, X. (2016). Vocabulary Acquisition: What does It Mean to Know a Word? *Studies in English Language Teaching*. Vol. 4. No. 1. 2016
- Zimmerman, C.B. (1997). *Historical Trends in Second Language Vocabulary Instruction*. Cambridge: Cambridge University Press.
- Vold, Martine (2018). *Factors that Predict Vocabulary and Self-Assessment in English as a Foreign Language*. MA thesis. University of Stavanger.

8. Appendix

8.1 Competence aims

Competence aims after Vg1 – programmes for general studies and Vg2 – vocational education programmes

Language learning

The aims of the training are to enable the apprentice to

- evaluate and use different situations, working methods and learning strategies to further develop one's English-language skills
- evaluate own progress in learning English
- evaluate different digital resources and other aids critically and independently, and use them in own language learning

Oral communication

The aims of the training are to enable the apprentice to

- evaluate and use suitable listening and speaking strategies adapted for the purpose and the situation
- understand and use a wide general vocabulary and an academic vocabulary related to his/her own education programme
- understand the main content and details of different types of oral texts about general and academic topics related to one's education programme
- listen to and understand social and geographic variations of English from authentic situations
- express oneself fluently and coherently in a detailed and precise manner suited to the purpose and situation
- introduce, maintain and terminate conversations and discussions about general and academic topics related to one's education programme
- use patterns for pronunciation, intonation, word inflection and various types of sentences in communication
- interpret and use technical and mathematical information in communication

Written communication

The aims of the training are to enable the apprentice to

- evaluate and use suitable reading and writing strategies adapted for the purpose and type of text
- understand and use an extensive general vocabulary and an academic vocabulary related to one's education programme
- understand the main content and details in texts of varying length about different topics
- read to acquire knowledge in a particular subject from one's education programme
- use own notes to write texts related to one's education programme
- write different types of texts with structure and coherence suited to the purpose and situation
- use patterns for orthography, word inflection and varied sentence and text construction to produce texts
- produce different kinds of texts suited to formal digital requirements for different digital media
- evaluate different sources and use contents from sources in an independent, critical and verifiable manner

Culture, society and literature

The aims of the training are to enable the apprentice to

- discuss and elaborate on culture and social conditions in several English-speaking countries
- present and discuss current news items from English language sources
- discuss and elaborate on the growth of English as a universal language
- discuss and elaborate on different types of English language literary texts from different parts of the world
- discuss and elaborate on English language films and other forms of cultural expressions from different media
- discuss and elaborate on texts by and about indigenous peoples in English-speaking countries
- select an in-depth study topic within one's education programme and present this

8.2 Request for Participation in the Research Project

Vil du delta i forskningsprosjektet

” English Vocabulary Size in upper secondary school”?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å måle vokabularstørrelsen hos elevene i videregående skolen. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål

Formålet med prosjektet er å undersøke elevenes vokabularstørrelse. Du som deltaker vil kunne få opplysning om hvor mange ord du kan på engelsk. Det er viktig for din forståelse når du leser, snakker, lytter eller skriver. For å kunne måle vokabularstørrelsen, skal du ta en enkel online test, samt en av dine skriftlige innleveringer skal analyseres.

Poenget med å gjennomføre dette prosjektet er å måle hvor mange ord kan elevene i videregående skolen, sammenligne det med resultater for lavere trinn, samt sammenligne med antall ord man skal kunne for å kunne gjennomføre ulike aktiviteter på engelsk (for eksempel lese en bok eller se en TV-serie).

Dette er en del av et masterstudium i lesevitenskap ved Universitetet i Stavanger.

Hvem er ansvarlig for forskningsprosjektet?

Universitetet i Stavanger er ansvarlig for prosjektet.

Hvorfor får du spørsmål om å delta?

Du blir spurt om å delta siden du er en elev ved skolen hvor undertegnede jobber.

Hva innebærer det for deg å delta?

- *«Hvis du velger å delta i prosjektet, innebærer det at du gjennomføre en online test. Det vil ta deg ca. 20-25 minutter. Spørreskjemaet inneholder 144 forskjellige ord, som du skal matche deres definisjon med. Dine svar fra spørreskjemaet blir registrert elektronisk.*
- *I tillegg, skal en av dine innleverte tekster analyseres med hensyn til vokabularet (hvor mange ord som er blitt brukt, hvor mange repetisjoner, hvor mange ord fra den Akademiske Ordlista er brukt mm.).*

Det er frivillig å delta

Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykke tilbake uten å oppgi noen grunn. Alle opplysninger om deg vil da bli anonymisert. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg. Det vil på ingen måte påvirke ditt forhold til skolen/læreren din.

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket. *Navnet ditt vil jeg erstatte med en kode som lagres på egen navneliste adskilt fra øvrige data. Deltakerne vil ikke kunne gjenkjennes i publikasjonen.*

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?

Prosjektet skal etter planen avsluttes 03.05.2019. Ved prosjektslutt skal alle opplysningene destrueres.

Dine rettigheter

Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke personopplysninger som er registrert om deg.
- å få rettet personopplysninger om deg.
- få slettet personopplysninger om deg.
- få utlevert en kopi av dine personopplysninger (dataportabilitet), og
- å sende klage til personvernombudet eller Datatilsynet om behandlingen av dine personopplysninger.

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra Universitetet i Stavanger har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Hvor kan jeg finne ut mer?

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- *Universitetet i Stavanger ved Torill Irene Hestetræet, tlf. 51831358, fakultetet for utdanningsvitenskap og humaniora.*
- NSD – Norsk senter for forskningsdata AS, på epost (personvernombudet@nsd.no) eller telefon: 55 58 21 17.

Med vennlig hilsen

Prosjektansvarlig
(Forsker/veileder)

Student

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet og har fått anledning til å stille spørsmål. Jeg samtykker til:

- å delta i *en online test*
- at lærer kan bruke mitt skriftlig arbeid i prosjektet*

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet, ca. *juni 2019*.

(Signert av prosjektdeltaker. dato)