Dissertations in Language Education

Stockholm University

English academic vocabulary knowledge among Swedish upper secondary school students

Marcus Warnby

Doctoral Thesis in Language Education at Stockholm University, Sweden 2023

English academic vocabulary knowledge among Swedish upper secondary school students

Marcus Warnby

Academic dissertation for the Degree of Doctor of Philosophy in Language Education at Stockholm University to be publicly defended on Friday 31 March 2023 at 13.00 in sal G, Arrheniuslaboratorierna, Svante Arrhenius väg 20 C.

Abstract

This thesis explores Swedish upper secondary school students' written receptive English academic word knowledge. Academic vocabulary are words that are more frequent in academic than in general discourse without being discipline-specific but frequent across disciplines, for example, *however*, *related*, *partially* and *delineate*. All the participants belong to study programmes which, according to the curriculum goals (Skolverket, 2013), prepare them for university studies, where English is a common reading language. Despite the university-preparatory goal, the syllabi contain no guidelines about academic English nor academic vocabulary.

The thesis is based on two premises: 1) academic vocabulary is a central component of reading at university, and 2) the curriculum goal of being prepared for university studies presupposes the ability to read literature in English.

The present thesis uses existing and validated tests targeting academic lexis. There are three validity arguments for using academic vocabulary measurements as indicators of students' predicted academic reading comprehension. First, because reading comprehension largely depends on word knowledge (e.g., Gough & Tunmer, 1986; Grabe & Stoller, 2019), measures of academic word knowledge inform about an essential component of academic reading comprehension. Second, to comprehend an academic text a reader should know 98 percent of the words in a text (e.g., Nation, 2001). In academic texts, approximately 10–14 percent of the words are academic. Thus, without a high degree of academic word knowledge, the 98 percent threshold cannot be reached. Third, if basic word knowledge is lacking, more nuanced knowledge aspects which may be important for academic deep reading are likely lacking too. For this reason, it is worthwhile testing a basic aspect of word knowledge first; to this end, this thesis tests the connection of a word form to its most common meaning.

Based on these premises and validity arguments, the thesis seeks to estimate the academic vocabulary knowledge of students at the beginning and the end of mandatory English instruction. Furthermore, factors that may explain this word knowledge are explored.

The thesis adopts a cross-sectional design where almost 1,000 participants were administered vocabulary tests, questionnaires and a survey of out-of-school English activities. Mainly statistical analyses were used.

The results reveal large variations in academic vocabulary knowledge within and between samples. Significant factors positively related to academic vocabulary are involvement with out-of-school English, age, gender and study programmes. Approximately half of the students leaving mandatory English courses do not reach the minimum threshold score indicating mastery of academic lexis.

Since there are no guidelines in the English syllabi about academic vocabulary knowledge, the outcomes are expected, namely large variations in and, on average, a low level of academic word knowledge. There is a risk that many students falling below the threshold are not sufficiently prepared for taking on university reading tasks. The disruption in constructive alignment between the curriculum goal and the syllabi guidelines should be considered and the thesis suggests a curriculum change where the English mandatory courses for university-preparatory programmes include principled instruction about academic English reading ability of which academic vocabulary knowledge is one central component.

Keywords: *Receptive English academic vocabulary knowledge, upper secondary school students, learners of English for academic purposes, academic reading literacy, Swedish curriculum, university-preparatory education.*

Stockholm 2023
http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-213970

ISBN 978-91-8014-178-9
ISBN 978-91-8014-179-6

Stockholm University

Department of Teaching and Learning

Stockholm University, 106 91 Stockholm

ENGLISH ACADEMIC VOCABULARY KNOWLEDGE AMONG SWEDISH UPPER SECONDARY SCHOOL STUDENTS

Marcus Warnby

Stockholm
University

English academic vocabulary knowledge among Swedish upper secondary school students

Marcus Warnby

ISBN print 978-91-8014-178-9
ISBN PDF 978-91-8014-179-6

Cover illustration: Linn Warnby, "It's abstract", 2023

This research was carried out within the SEMLA research school "Learning, teaching and assessment of Swedish and English. Multilingualism as an asset and a challenge". The SEMLA is a PhD programme for teacher educators, funded by Vetenskapsrådet (VR 2017:06048) and coordinated by Stockholm University in collaboration with the University of Gothenburg and Umeå University.

Printed in Sweden by Universitetsservice US-AB, Stockholm 2023

Preface and acknowledgements

In many ways, I feel I have grown during these recent four and a half years. There are several events and many people who have been important for this growth.

I have grown as a professional. My previous higher education and diplomas including a Master of Arts in the subjects of French and Swedish, an upper secondary school teaching certificate, a Master in Applied Linguistics and an advanced level certificate as a Swedish school principal, together with many years working as an upper secondary school teacher, a junior lecturer at the teacher education departments in both Stockholm and Gothenburg, a lower secondary school principal and as a test developer for the national test system for foreign languages. Together, these experiences have certainly led me to this stage in some ways but the experience of taking on a PhD has been an altogether different experience.

My network of people has grown, and my existing networks have grown stronger. First and foremost, an important new network for this thesis has been the schools which have participated, my thanks to all those students who have willingly participated and to all the teachers and school principals who generously agreed to be part of this research. Without you, this thesis would never exist. Two other people have earned my deepest gratitude, namely my supervisors Camilla Bardel and Hans Malmström. Thanks for your guidance and support through both the highs and the lows of my research journey. Thanks for putting up with me and for guiding me in narrowing down all the hundreds of research ideas that I have proposed during our zoom-meetings, our e-mail and SMS conversations. My heartfelt thanks also go to Kajsa Yang Hansen who joined later in the process as an additional supervisor. A special thank you goes to Philip Shaw and Henrik Gyllstad for thoroughly reading, discussing and giving feedback on my thesis at previous stages. I am also thankful to Professor James Milton for consenting to read and discuss this thesis at the public defence. Thank you to Una Cunningham and Viveca Lindberg for the final review. Thanks to all the members of the FUR-research group (University of Gothenburg) chaired by Monica Rosén, Joanna Giota and Ulrika Wolff; being a part of your positive and critical research group is inspiring, thought-provoking and educational; you all know who you are. Thanks, Erika Majoros, you R a coding expert. Thanks to Nicole Busby in Norway and Linda Eriksson in Örebro for the exchange of ideas regarding academic English; to Zakaria

Lemmouh and Per Snoder for our interesting discussions on vocabulary research. Many thanks, too, to my dear friends Per Blomqvist and Göran Söderlund for the encouragement, support and laughs during the years; our interest in improving research and education seems to grow whenever we talk. Thanks to Girma Berhanu for pushing me to apply for a doctoral position. Thanks to the department at Stockholm University, all my present and former doctoral student colleagues, my Italian room mate Francesco Vallerossa, to our Director of studies Tore Nilsson and our Head of department Anna Chryssafis. A most precious network is the SEMLA research school doctorates (most of which have already achieved PhDs): Sara, Jasmine, Ingela, Erika, Elisabeth O, Elisabeth N, Clara and Anna-Marie; thanks for all the shared laughter, tears and anger, for all the intellectual and non-intellectual conversations. I am sure we will continue growing together as a group. Thank you, Margareta, for all your administrative work within the research school. Thanks to the all the people involved in the SEMLA. I am very thankful to the SEMLA steering group: Camilla, Gudrun, Liss Kerstin and Eva for informed comments during these years and, mostly, for granting me this growth within the research school.

Furthermore, my academic English vocabulary has grown. Writing a thesis in a language and about a language that I have not studied for around 30 years certainly has been a challenge. Janet Enever, thank you for helping me improve my English academic writing skills, and, also for inspiring discussions about educational equity and equality from the beginning of the SEMLA.

Also, it has taken much time, despair and hair loss to manage to solve the analysis of a relatively large data collection, redesign the thesis during school lockdowns, understand how to solve statistical data analyses and remain calm during lengthy journal review processes. In more positive words, one could say that I have ‘grown a forehead’.

Moreover, as a child, I grew up with my mother and father’s unconditional love and support. Barbro and Svante, you are the best parents.

Finally, the most important people in my life are my beloved and crazy family: my children Linn, Sixten and Charlie and my wife Åsa. Thanks for being there all the time, believing in me and making my world a more beautiful and happy place. My love for you grows stronger every day, forever and ever.

Torslanda, February 2023
Marcus

Included studies

The following three studies constitute the basis for the present thesis:

I. Warnby, M., Malmström, H., & Yang Hansen, K. (2023). Linking scores from two English academic vocabulary tests – the VLT-Ac and the AVT. *Language Testing*, Advance Online Publication, 12 January, 2023. https://doi.org/10.1177/02655322221145643

II. Warnby, M. (2023). Academic vocabulary knowledge among adolescents in university preparatory programmes. *Journal of English for Academic Purposes*, *61*(2023) 101203, 1–14. https://doi.org/10.1016/j.jeap.2022.101203

III. Warnby, M. (2022). Receptive academic vocabulary knowledge and extramural English involvement – Is there a correlation? *ITL – International Journal of Applied Linguistics*, *173*(1), 120-152. https://doi.org/10.1075/itl.21021.war

Contents

List of tables

List of figures

Abbreviations

AVL	The Academic Vocabulary List
AVT	The Academic Vocabulary Test
AWL	The Academic Word List
BNC	The British National Corpus
CEFR	The Common European Framework of Reference for Languages: Learning, teaching, assessment
COCA	The Corpus of Contemporary American English
CTT	Classical Test Theory
EE	Extramural English
EFL	English as a foreign language
EK	Business Management and Economics Programme (Ekonomi-programmet)
ES	Arts Programme (Estetiska programmet)
FL	Foreign language
HU	Humanities Programme (Humanistiska programmet)
IELTS	International English Language Testing System
IRT	Item Response Theory
L1	First language
L2	Second language
NA	Natural Science Programme (Naturvetenskapsprogrammet)
NAE	The Swedish National Agency for Education (Skolverket)
SA	Social Science Programme (Samhällsvetenskapsprogrammet)
TE	Technology Programme (Teknikprogrammet)
TP1	First Time Period of data collection, entry-sample
TP2	Second Time Period of data collection, exit-sample
VKS	Vocabulary Knowledge Scale
VLT	The Vocabulary Levels Test
VLT-Ac	The academic section of the Vocabulary Levels Test
VST	The Vocabulary Size Test

1. Introduction

> It is the responsibility of the school that *all individual students* [emphasis added] … in a national higher education preparatory programme in the upper secondary school are given the opportunity to fulfil the requirements for a diploma providing eligibility for studies in higher education, which means that the student *has sufficient knowledge to be well prepared for studies in higher education* [emphasis added].
>
> – *Curriculum for the upper secondary school* (Skolverket, 2013, Ch. 2.1. Knowledge – Goals, p. 8)

The Swedish national curriculum for upper secondary education contains overarching and abstract norms, values and goals to be reflected in, and attained with, the instruction. The curriculum citation above shows a central objective for the upper secondary programmes providing eligibility to higher education, namely that each individual student develops the basic knowledge required to study at tertiary level. This thesis assumes that any student continuing to higher education will need a certain level of English academic vocabulary knowledge to be well prepared for the English academic reading demands in Swedish universities. In the thesis, Swedish upper secondary school students' written receptive English academic vocabulary knowledge will be assessed, explored and discussed. The thesis' focus is thus one facet of the above-cited educational objective.

1.1. Background

Words are essential for conveying meaning in any language (Lewis, 1993; Milton, 2009; Schmitt, 2010; Wilkins, 1972) and, in academic language, a key component is academic vocabulary (Coxhead, 2016; Nagy & Townsend, 2012). Broadly speaking, words in academic discourse can be categorised into general, academic and disciplinary vocabulary (Chung & Nation, 2003; Coxhead, 2016; Nation, 2013; Pecorari et al., 2019; Skjelde & Coxhead, 2020). Normally, *academic vocabulary* is defined as the words typical for academic texts and common across disciplines (Charles & Pecorari, 2015; Coxhead, 2000, 2016; Gardner & Davies, 2014). At university, advanced Eng-

lish proficiency has become a requirement due to the international use of English for academic purposes (e.g., Malmström & Pecorari, 2022; Mauranen et al., 2016; Salö, 2016). Consequently, knowledge of English academic vocabulary is indispensable for any student regardless of his/her first language (L1) in order to read and understand texts in much internationalised higher education. However, without specifically addressing academic vocabulary, university students often find the academic language difficult to understand (Ambjörnsdóttir & Ingvarsdóttir, 2018; Hellekjær, 2009; Leese, 2010), are overwhelmed by the large quantity of course reading materials (Perander et al., 2020) and state lack of reading habits to be a reason for difficulties in academic tasks (Alharbi, 2017). Students report that vocabulary constitutes "the hardest part of reading in English" (Delgadillo-Collazos, 2020, p. 222). Similarly, Norwegian upper secondary school students who are under-achieving in academic reading comprehension tasks state that their unfamiliarity with many words is the principal impediment to comprehension (Hellekjær, 2009). Regarding academic vocabulary knowledge, research suggests that many Norwegian (Skjelde & Coxhead, 2020) as well as Icelandic (Edgarsson, 2018) upper secondary school students preparing for university have limited English academic vocabulary knowledge. A similar situation is evident in the context of this thesis project – Sweden.

The majority of students entering a Swedish university will have to cope with literature published in English already at the undergraduate level (Kuteeva, 2014; Malmström & Pecorari, 2022; Pecorari et al., 2011). Swedish university students themselves have reported negative emotions regarding reading English texts, often due to the large number of unknown words (Eriksson, 2021, 2022; Pecorari et al., 2012). This recalls the anecdotally reported issue of students' preparedness for academic studies, a debate which centers around the poor level of academic literacy with which students from upper secondary school enter university (Enefalk et al., 2013; Frankki, 2015; Samuelsson, 2013; Svedin, 2017).

For admission to Swedish higher education, students are required to have a pass grade in the final upper secondary mandatory English course, English 6, which is considered equivalent to the B2-level of the *Common European Framework of Reference for Languages: Learning, teaching, assessment* (CEFR) (Skolverket, 2021a). There are six national upper secondary programmes providing eligibility to higher education and they all share the objective of providing their students with equitable and good preparation for tertiary level studies. In fact, the overarching name of these programmes is *higher education preparatory programme* (Skolverket, 2014) and it is the school's responsibility to ensure that all individual students have "sufficient knowledge to be well prepared for studies in higher education" (Skolverket, 2013, p. 8). However, there is limited research on the academic language knowledge among this group of English-as-a-foreign language (EFL) learners preparing for university.

Despite the importance of academic English vocabulary knowledge for efficient academic communication already at the onset of tertiary level studies, very little is mentioned about English academic language in the upper secondary curriculum. In fact, the English syllabi do not offer any principled approach or concretely expressed content of instruction regarding academic vocabulary. Apparently, it is assumed that academic vocabulary will be learned incidentally, as a by-product of other meaning-oriented activities, inside or outside school.

For each upper secondary programme, there are programme-specific diploma goals. For example, in the Technology Programme, the diploma goals state explicitly that students must learn English to be able to acquire knowledge from current technological developments. Since the programmes have differently expressed diploma goals, it may be worthwhile to use the study programme as a factor that may explain difference in English academic vocabulary knowledge. In this connection, prior English grades must be taken into account since students enter programmes with different English proficiencies. Also, academic vocabulary knowledge can explain achievement (Masrai & Milton, 2018; Townsend et al., 2012), and, therefore, it could be worthwhile to correlate academic vocabulary knowledge with upper secondary English final grades.

From previous research, it is known that general vocabulary may be incidentally learned through extensive out-of-school, so-called extramural English exposure (e.g., Sundqvist, 2009; Sundqvist & Sylvén, 2016; Sylvén, 2006b). Furthermore, differences in general vocabulary knowledge between males and females have been explained as an effect of gender group differences in out-of-school English activities (e.g., Peters et al., 2019; Sundqvist & Wikström, 2015). Whether there are similar patterns regarding academic vocabulary knowledge may therefore be interesting to explore.

Moreover, socioeconomic status (SES) may affect language development. For example, the typical pattern is that children from a low SES background display lower receptive vocabulary size and lower language ability in comparison to more affluent children whether it is in L1 or in second language (L2) (Calvo & Bialystok, 2014). Previous research presents evidence that parental education correlates highly with general reading ability in L1 (e.g., Myrberg & Rosén, 2009) as well as in L2 (e.g., Abbasian et al., 2020). Since vocabulary knowledge is a good predictor of reading comprehension, one assumption could therefore be that students' parental educational level may influence students' L2 English academic vocabulary knowledge.

A further factor to be considered may be the proposal that individuals with more than one L1 can profit from their larger language repertoire during third language acquisition, compared to those with only one L1 (Bartolotti & Marian, 2017). Sweden is a multilingual country with circa 200 different spoken languages (Institutet för språk och folkminnen, 2021) and nearly 20 percent

of the population was born abroad (Statistics Sweden, 2022). Considering current research interest and discussion on the role of multilingualism for additional language learning (see, e.g., Bartolotti & Marian, 2017; Bonnet & Siemund, 2018; Fuster, 2022) it is reasonable to investigate whether a Swedish EFL learner with more than one L1 has a larger English academic vocabulary size than a peer with one L1[1].

Although there are patterns of unequal recruitment to universities related to factors such as, for example, gender, SES and age (Swedish Council for Higher Education, 2016) there are conflicting results as to what factors explain tertiary level study drop-out (Statistics Sweden, 2012; Swedish Council for Higher Education, 2017). The link between such factors and academic literacy as measured with academic vocabulary knowledge among pre-university students could provide more knowledge to the field.

In sum, very little is known about the proficiency level of written receptive L2 English academic vocabulary knowledge among students in university-preparatory programmes and whether this knowledge can be explained by, for example, the above-mentioned educational and other individual factors.

1.2. Premises, research questions and aims

Building on the background setting outlined in the previous section, the purpose of this thesis is to further our knowledge about the written receptive English academic vocabulary knowledge among students in upper secondary programmes aiming to prepare them for tertiary level studies. The thesis is built on two premises. The first premise is that academic vocabulary knowledge is a central facet of academic reading comprehension in university studies. The second premise is a consequence of the first: that academic vocabulary knowledge is related to the curriculum objective of providing all students with sufficient knowledge to be well prepared for studies in higher education. These two premises are the basis for the following overarching research questions:

[1] It is beyond the scope of this thesis to deconstruct the concepts of multilingualism and language repertoires. Acknowledging that the concept *language* is organic and that language is construed in sociocultural settings, in this thesis, however, languages are described as separate linguistic systems (e.g., English or Swedish) which contain different varieties and registers (e.g., academic English or colloquial English). Learners may draw on all their resources to acquire new knowledge and, thus, add to their language repertoire. Recent research about multilingualism and language repertoires within the Swedish school system suggests that research may investigate, for example, the concept of language repertoires in relation to a person's languages expressed as separate entities (e.g., Bylund, 2022; Snoder, 2022).

RQ1: What is the written receptive English academic vocabulary knowledge among Swedish upper secondary school EFL learners?

RQ2: To what degree can specific educational and other individual factors explain the written receptive English academic vocabulary knowledge among Swedish upper secondary school EFL learners?

The factors in focus in this study are (*i*) English exit grades from compulsory school and upper secondary school (mandatory instruction), (*ii*) study programme, (*iii*) gender, (*iv*) number of years of formal EFL instruction, (*v*) number of L1s, (*vi*) age, (*vii*) parental educational level and (*viii*) students' amount of extramural English involvement.

To answer and discuss these overarching research questions, the thesis builds on three studies which aim to:

a) identify a measure of the construct *written receptive English academic vocabulary knowledge* capable of yielding valid usage of scores pertaining to the construct and the target population,

b) map students' written receptive English academic vocabulary knowledge at the beginning and end of the mandatory part of English instruction in upper secondary English school, and

c) explore potential factors of importance for written receptive English academic vocabulary knowledge.

1.3. Studies I – III overview

The thesis comprises three studies gathering data from a total of 998 consenting participants. Each study has its separate focus addressing one of the three aims (see section 1.2.) but all three studies add, to a varying extent, information integral to answering the thesis' overarching research questions.

Study I (Warnby et al., 2023) investigates the links and valid uses of two measurement instruments of academic lexis within the target population. The study explored different Item Response Theory (IRT) models for the purpose of linking 385 upper secondary school students' scores on two tests of written receptive English academic vocabulary addressing the aspect of meaning-recognition taking partial lexical knowledge into account – the academic section in the revised Vocabulary Levels Test (VLT-Ac; Schmitt et al., 2001) and the Academic Vocabulary Test (AVT; Pecorari et al., 2019). These tests target two different operationalisations of academic vocabulary, the former with academic words grouped as word families (The Academic Word List (AWL; Coxhead, 2000)), and the latter as lemmas (The Academic Vocabulary List (AVL; Gardner & Davies, 2014)). Study I is largely connected to the first aim of the thesis. However, it is also connected to the second aim since it reports

on certain measures of written receptive English academic vocabulary knowledge.

Study II (Warnby, 2023) speaks mainly to the second aim by mapping academic vocabulary knowledge among students entering upper secondary programmes and students ending their final mandatory English course. By using cross-sectional data from 952 participants, the study estimates the written receptive English academic vocabulary knowledge and development at the beginning of and the end of mandatory upper secondary English instruction. Furthermore, final grades in English and AVT scores are correlated, and logistic regression analyses are used to explore the likelihood for different student groups to reach suggested academic vocabulary mastery thresholds. Study II is also, to a certain extent, connected to the third aim of the thesis since English school grades, gender and study programmes are explored as factors that may be related to the development of written receptive English academic vocabulary knowledge.

Study III (Warnby, 2022), which mainly addresses the third aim of the thesis, correlates several factors with academic vocabulary knowledge among a combined sample of 817 students. Although the focus is on how extramural English (EE) may be a source of incidental learning of academic vocabulary, the study uses additional factors (number of years of formal EFL instruction, age, gender, parental educational level, number of L1s) in a regression analysis to understand what factors may explain the variance in academic vocabulary knowledge. Furthermore, since, for example, the internal consistency and discrimination indices in the AVT are analysed, study III adds information about the valid and reliable use of the AVT, thus connecting to the first aim.

1.4. Reading guide and thesis outline

This thesis should be understood from a language education perspective, broadly speaking. Its contribution connects primarily to educational assessment with a specific focus on written receptive English academic vocabulary knowledge connected to the upper secondary educational goal of university preparation. Importantly, this thesis does not add to knowledge regarding how academic words are efficiently taught or how to best assess the full construct of academic reading. Instead, the thesis draws on other specialists' work within such research fields to be able to present a case about Swedish education in relation to the thesis' empirical findings.

In the current chapter, background, premises, research questions and aims have been established. Chapter 2 provides a more detailed contextualisation of the thesis with respect to the position of English in Sweden and the Swedish education system. Chapter 3 presents theoretical concepts regarding vocabulary in general. Chapter 4 offers an account of the main construct in this thesis: academic vocabulary. A review of previous research on academic vocabulary

related to the thesis topic is presented in Chapter 5. In Chapter 6, the research design and an overview of the data collection process are presented, including descriptions of the instruments used. In connection with these presentations, critical issues pertaining to the validity and reliability of the instruments and analyses are discussed. The chapter ends with some ethical considerations. Chapter 7 summarises the main findings of the three studies. Finally, in Chapter 8, the findings from the three studies are discussed in relation to the thesis' premises and research questions as well as to educational implications. Furthermore, Chapter 8 will acknowledge some of the limitations of the thesis and suggest future research. Chapter 9 offers a summary of the thesis in Swedish.

2. The Swedish study context

This chapter situates the project with respect to the place and role of English in Sweden. Moreover, a brief overview of the Swedish education system is given mainly with attention to the upper secondary education, English as curricular subject and English at Swedish universities.

2.1. English in Sweden

Although Swedish is the main language in Sweden as stipulated in the Language Act (SFS 2009:600), the 10 million inhabitants contribute to an immensely diverse linguistic society (European Commission, 2018). In addition to Swedish, the three most frequent L1s are Arabic, Finnish, and Bosnian-Croatian-Serbian (Parkvall, 2016, 2019; Språkrådet, 2022). While English is not part of the 10 most frequent L1s, the general English ability is high, and Swedish adolescents are repeatedly placed among the top proficiency users of English as a second language (L2) or a foreign language (FL) internationally (Bonnet, 2004; Education First, 2019, 2021; Erickson, 2004; European Commission, 2012; Simensen 2010).

The English language is widely used in Swedish society as in many other countries; English may be found in advertisements and branding, on TV and in movies (where subtitling is more prevalent than dubbing), as the language of communication within companies, and so forth (e.g., Enever, 2018; Sundqvist, 2019; Sylvén, 2019). Furthermore, the shift in the digital landscape over the past two decades with children's and adolescents' steadily increased use of, for example, social media, online gaming and streamed movies/clips (Statens Medieråd, 2019) has dramatically widened the opportunities to encounter and acquire English. As a consequence of Swedish adolescents' increased exposure to and involvement with out-of-school English the concept of extramural English (EE) has been coined in the Swedish research context. The concept was first presented by Sylvén (Sylvén, 2006a, 2006b), thoroughly investigated by Sundqvist (2009), and, thereafter, theoretically broadened by the two scholars (Sundqvist & Sylvén, 2016). Several studies on EE have revealed not only that adolescents to a large extent are engaged in English-mediated activities on a daily basis but also that this EE involvement is beneficial

for the development of general English ability and general vocabulary knowledge (Olsson, 2016; Skolverket, 2012a; Sylvén & Sundqvist, 2012).

In relation to the strong position of English in Sweden, Swedish children, adolescents and adults have a strong belief in the relevance of being proficient in English and therefore may rather invest in English than any other foreign language or may express ideas of English being a sufficient L2 in addition to one's mother tongue (Bardel et al., 2019; Bylund, 2022; Finndahl, in progress; Henry, 2012). In view of the strong position of English in Sweden, debates have emerged about the status of English as an L2 or an FL (e.g., Falk et al., 2015; Hult, 2012; Hyltenstam, 2004; Phillipson, 1992; Sundqvist & Sylvén, 2016). It is beyond the scope of this thesis to settle this debate, and the notions of English as an FL (EFL) and English as an L2 will be used interchangeably when referring to the Swedish context.

2.2. English in Swedish education

This section first provides some general information about the Swedish school system in sub-section 2.2.1 and some specific information about upper secondary education in sub-section 2.2.2. Thereafter, the role and place of English in upper secondary school (section 2.2.3) and at universities (2.2.4) are presented. The selected description of English within the Swedish curriculum below is based on its relevance to the thesis. For the interested reader, a recent and more detailed evaluation of teaching English in secondary school in Sweden can be found in Siegel (2022) (see especially chapter 3, "Examining steering documents: From CLT to CEFR to Skolverket").

2.2.1. Essentials about the Swedish school system

Within the Swedish educational system a distinction can be made between school, including preschool, primary, lower secondary and upper secondary school, and post-upper secondary education, including, for instance, higher vocational education, university college and university. Compulsory schooling starts with one year in preschool class when children turn six, and ends ten years later when they are in school year 9 (i.e., the 10th year of schooling). Thereafter, most 15-year-old students continue to three years of non-mandatory upper secondary school, which offers a variety of national programmes, vocational as well as university-preparatory[2].

[2] The number of students entering a national programme differs from year to year. During the school year 2021/2022, 111,392 out of a total of 133,261 students were enrolled in a national programme, which converts to 83.59% (Skolverket, 2022d).

Regardless of school form, the school system is based on a common set of norms and values, and epistemological and ontological assumptions. For instance, in the Swedish Education Act (SFS 2010:800) it is stated that education within the school system should be based on principles of equality, regardless of who you are, where you come from, or where your school is located:

> I utbildningen ska hänsyn tas till barns och elevers olika behov. Barn och elever ska ges stöd och stimulans så att de utvecklas så långt som möjligt. En strävan ska vara att uppväga skillnader i barnens och elevernas förutsättningar att tillgodogöra sig utbildningen. …
> Utbildningen inom skolväsendet ska vara likvärdig inom varje skolform … oavsett var i landet den anordnas.
>
> [Children's and students' different needs must be considered within the education. Children and students must be given support and stimulation so that they develop to the greatest extent. An ambition must be to counterbalance the differences in children's and students' conditions of absorbing the education. …
> The education within the school system must be equivalent within every type of school … regardless of where in the country it is organised.] (SFS 2010:800, 1 Ch, 4§ & 9§, my translation)

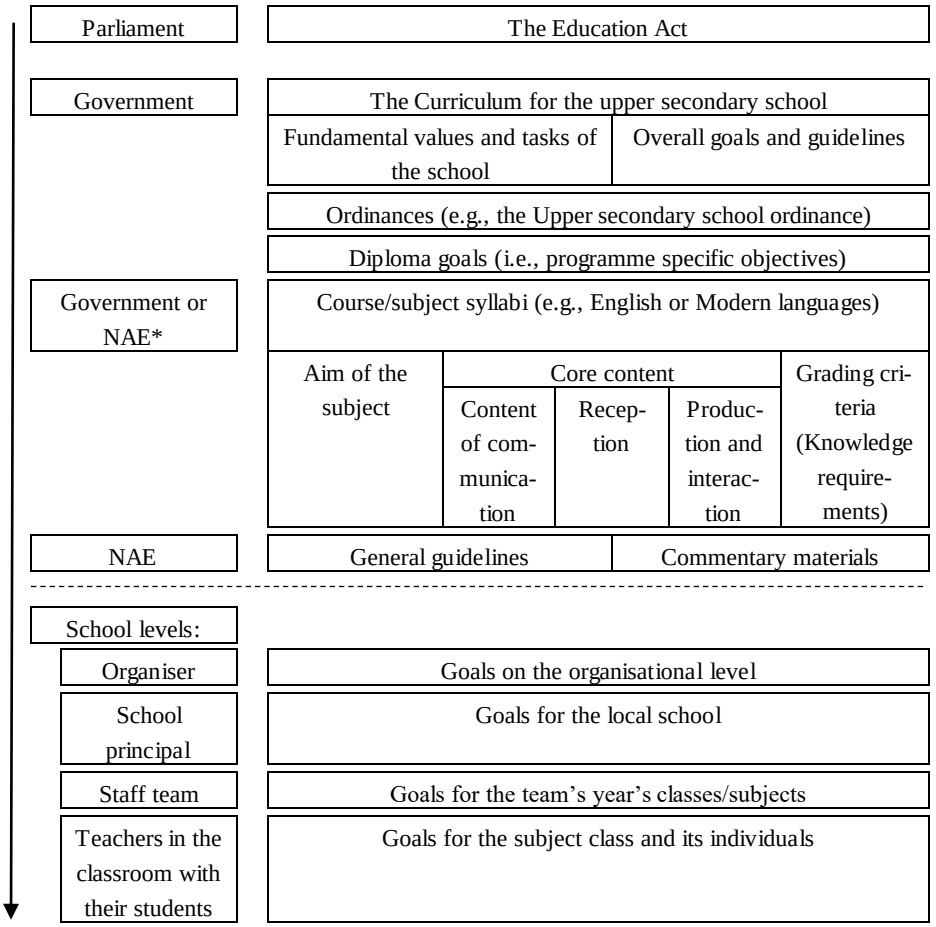
The Education Act provides the principles for the curriculum, which in turn provides a framework for the ordinances and the diploma goals. The curriculum contains two sections: 1) Fundamental values and tasks of the school and 2) Overall goals and guidelines. Both values and goals are set by the government and are all superordinate to the subject courses' syllabi. The syllabi are set by the Swedish national agency for education (NAE) except for the syllabi for nine subjects common to all upper secondary programmes which are set by the government. Among those common subject disciplines are English, Swedish, history, mathematics, physical education and social science. However, the NAE has a role to play in formulating these syllabi which are then proposed for governmental approval. The syllabi are divided into three sections: 1) Aim, 2) Core content of instruction and 3) Knowledge requirements[3] (Regeringen, Dir. 1991:117).

[3] A new curriculum, Lgy22, was implemented in July 2021. As this thesis collected data during the time of the previous curriculum, Gy11, it references that curriculum. For a recent review of the new English syllabi and the changes, see Siegel (2022). The Swedish national agency for education offers a document with comparisons between the former and the new English syllabi (Skolverket, 2022b). One of the changes is the return of the term *grading criteria* (Swe. *Betygskriterier*) used in the curriculum Lgy94, from 1994 to 2011, thus replacing the term *knowledge requirements* (Swe. *kunskapskrav*) in the previous curriculum, Gy11.

All these officially stated steering documents for education are interpreted and enacted at the local level (school organisers, school principals, teacher teams, classroom teachers and students). Figure 1 illustrates what, from a macro (national level) to a micro (local/classroom) level, the *national curriculum* looks like in the Swedish school system.

Figure 1

The Swedish school system's actors and steering documents – the case of upper secondary education.



* The Swedish National Agency for Education (NAE) sets all subject syllabi except for the syllabi for the common subjects in upper secondary school programmes which are set by the government.

The term *curriculum* (Swe. *läroplan*) may have many meanings in an educational context. Often, in Sweden, it refers to all documents from the curriculum down to the syllabi mentioned in the system of steering documents above (e.g., Csöregh, 2022; Sturk, 2022). As such the term *curriculum* should be interpreted broadly, for example, when NAE states that "it is important to read the different parts of the curriculum as a whole in order to understand the purpose of the education" (Skolverket, 2018). However, *curriculum* may also be interpreted as the subject-specific document providing purpose and guidance to the overall instruction of the subject; as such, the document is positioned as superior to the subject courses' syllabi (Graves, 2016; Siegel, 2022). Moreover, adding complexity to *curriculum* as equal to nationally written policy tools, *curriculum* may theoretically be deconstructed into three parts that are sequentially ordered: 1) the *intended curriculum*, indicating what is to be taught and learned according to official policy tools, 2) the *enacted curriculum*, indicating what is happening in classrooms, and 3) the *learned curriculum*, indicating what the achievement and outcome is among the learners. In this thesis, the term *national curriculum* will be used when referring to the whole chain of national steering documents, while *curriculum* will mainly be used to refer to the specific document containing the norms, values and overall goals for the school and the terms *syllabus* and *English curriculum* will be used only for the subject-specific course document including the aim of the subject, the core content and the knowledge requirements. When required, *curriculum* will be specified, for example: the intended curriculum as opposed to the enacted curriculum.

Related to the curriculum issues above, a general curricular principle, worthy of mentioning since it is advocated for by the Swedish National Agency for Education (Skolverket, 2020), is the notion of *constructive alignment* (Biggs, 2003) which, broadly speaking, focuses on the relationship between educational objectives, teaching and learning outcomes. By adopting this principle, teachers deliberately align their interpretation of the curriculum intentions with the implemented teaching and learning activities and with the assessed learning outcomes by expressing, for example, clear and specific goals. The objective of constructive alignment is to achieve an effective educational design. The constructive alignment principle can also be argued to apply to the steering documents, in which the different levels of the national curriculum should coherently align intended overall goals with intended instructional content and the grading criteria.

2.2.2. Upper secondary education

In Swedish upper secondary education, there are 18 national programmes, comprising six university-preparatory programmes and 12 vocational programmes, all of which are three years in duration (Skolverket, 2012b). In order to be admitted to one of the 18 national programmes, students' final grades in school year nine are used. One of the criteria for admission is a pass grade in English, equivalent to CEFR-B1.

The six university-preparatory programmes (of particular interest in this thesis) are the Business Management and Economics Programme (Ekonomiprogrammet; EK[4]); the Arts Programme (Estetiska programmet; ES), the Humanities Programme (Humanistiska programmet; HU); the Natural Science Programme (Naturvetenskapsprogrammet; NA), the Social Science Programme (Samhällsvetenskapsprogrammet; SA) and the Technology Programme (Teknikprogrammet, TE).

In the curriculum for upper secondary school (Skolverket, 2013), overall objectives set out the knowledge that "all students should have acquired by the time they complete their schooling" (p. 8). Among these educational objectives, as referenced in section 1.1., it is stated that the responsibility of the school is that all individual students "in a national higher education preparatory programme in the upper secondary school are given the opportunity to fulfil the requirements for a diploma providing eligibility for studies in higher education, which means that the student has *sufficient knowledge to be well prepared for studies in higher education* [emphasis added]" (p. 8). Furthermore, the school is responsible for ensuring that each student acquires sufficiently good knowledge so that he/she "can use this knowledge for further studies" (p. 8).

The curriculum also stipulates that schools cooperate with, for example, universities, as this is required "for students to receive education of high quality, and serves as a basis … for further studies" (p. 12). Moreover, the "universities and university colleges … have important roles to play in providing information to schools and students" (p. 12) with respect to the societal landscape that constantly changes "in terms of the need for competence" (p. 12). It is very clear that the curriculum emphasises, on an abstract and overarching level, the importance of higher education within the university-preparatory programmes; however, Swedish research indicates that the language syllabi are poorly connected to academic literacy abilities (e.g., Eriksson et al., 2023; Palm, 2023; Warnby & Lemmouh, 2021).

[4] Henceforth, in this thesis, the Business Management and Economics programme is referred to as the Economics programme.

The six university-preparatory programmes offer broad eligibility for higher education; however, each programme has its own specific Diploma goals (Skolverket, 2022a). All six Diploma goals unanimously state that, after graduation, all students should have the knowledge resources required for higher education studies, mainly in relation to the programme profile areas. For example, in the Diploma goals for NA, it says:

> Efter examen från programmet ska eleverna ha kunskaper för högskolestudier inom främst naturvetenskap, matematik och teknik men även inom andra områden
>
> [After passing the diploma, students will have acquired knowledge for higher education studies, mainly within natural sciences, mathematics and technology but also within other areas.) (Skolverket, 2022a, para. 1, my translation)

In comparison, the Diploma goals for EK state the following;

> Efter examen från programmet ska eleverna ha kunskaper för högskolestudier inom främst ekonomi, juridik och andra samhällsvetenskapliga områden.
>
> [After passing the diploma, students will have acquired knowledge for higher education studies, mainly within economics, law and other social science areas.] (Skolverket, 2022a, para. 1, my translation)

In the diploma goals for two of the programmes, SA and TE, English is explicitly mentioned as an important language of communication within the study profile; in the diploma goals for SA it stated that:

> Utbildningen ska utveckla elevernas förmåga att skriva, läsa, tolka och förstå olika typer av texter inom utbildningens olika kunskapsområden. Eleverna ska också ges möjlighet att uttrycka sig i varierande skriv- och talsituationer på framför allt svenska och engelska.
>
> [The education shall develop the students' ability to write, read, interpret and understand different types of texts from the programmes different areas of knowledge. The students shall also be given the opportunity to express themselves, especially in Swedish and English, in a variety of written and oral contexts.] (Skolverket, 2022a, para. 7, my translation)

In the diploma goals for TE, it is stated that:

> Utbildningen ska ge eleverna kunskaper om och färdigheter i engelska i en teknisk kontext, så att de kan utveckla sin kommunikativa förmåga och därmed ta del av teknik och teknikutveckling.
>
> [The education shall provide the students with knowledge and skills in English in a technology context, so that they can develop their communicative ability and, therefore, follow and gain from technology and technology development.] (Skolverket, 2022a, para. 5, my translation)

2.2.3. English in the school system

English is the first foreign language to be taught in compulsory school and is a core subject in the Swedish school system on a par with Swedish/Swedish as a second language and mathematics. English instruction in Sweden starts in school year 3, at the latest (but mainly introduced from year 1). By year 9, school students will have received approximately 480 hours of EFL instruction during their compulsory (primary to lower secondary) schooling. In upper secondary education, the hours in the two mandatory English courses for students preparing for university – English 5 and English 6 – amount to approximately 180 hours of EFL instruction, provided during years 1 and 2 of the three year programme. This means that students preparing for university have one year off from English instruction even if they continue directly to university after school graduation. There is an optional English course in year three, English 7, but not all students take this course. In 2021, 62,535 third year students ended their upper secondary schooling in Sweden, of which 37,253 students had followed and completed English 7 with a pass grade (Skolverket, 2022d). It is not clear why not all students take English 7, however, Swedish adolescents' rather high general English proficiency and their engagement with English outside school may lead to their lack of enthusiasm and willingness to engage with more formal English learning within the school, "where students do not feel particularly challenged, tending to regard English lessons as providing a welcome opportunity for rest and relaxation" (Henry, 2014, p. 19).

The English syllabi for compulsory school and upper secondary education are interconnected and share many similarities. First, the overall aim expressed in the English subject syllabus for compulsory school and in the syllabi for the English courses in upper secondary school is similar. For instance, English instruction aims at developing students' *all-round communicative ability* (Swe. *allsidig kommunikativ förmåga*) (Skolverket, 2022c, 2022f) which indicates that the English curriculum is embedded in a communicative language teaching (CLT) approach (see section 3.5.). Second, for each sylla-

bus, the Swedish NAE provides commentary materials explaining and defining the syllabus' content and expressions in more detail. The commentary materials for the English syllabi (Skolverket, 2021a) describe how the content builds on and relates to the CEFR (Council of Europe, 2001). For instance, the English syllabi are said to conform to the communicative and action-oriented approach of the CEFR in that they define and emphasise the linguistic actions students are to develop rather than the building blocks needed for these linguistic actions (e.g., Siegel, 2022). With respect to this thesis' focus on vocabulary, it is, therefore, not surprising that vocabulary is more or less identified as absent in the syllabi (Bergström et al., 2022b; Siegel, 2022; Warnby & Lemmouh, 2021). In his description of the English curriculum, Siegel concludes that the two curriculum content sections focusing on language skills (receptive and productive) set out a number of different abilities to be developed by the students.

> These span all four main language skills, in line with a balanced CLT approach: listening, speaking, reading and writing. Explicit attention is also paid to language systems (i.e., pronunciation, intonation, and grammar), although the emphasis is clearly not on accuracy but rather comprehensibility; that is, making oneself understood. *Interestingly, the term vocabulary is not used* [emphasis added], although it could be implied from phrases like fixed expressions (e.g., *wake up, in my opinion, a long time ago*). (Siegel, 2022, p. 86)

Although vaguely addressed, it should be acknowledged that *vocabulary* has been introduced in the new English Lgy22 syllabi within the content section "production and interaction" – that is, not as a part of instruction for receptive skills – and merely as a facet of all kinds of linguistic phenomena used in productive and interactive skills:

> Språkliga företeelser, däribland uttal, vokabulär, grammatiska strukturer och meningsbyggnad, stavning, textbindning, inre och yttre struktur samt anpassning, i elevernas egen produktion och interaktion.
>
> [Linguistic phenomena, including pronunciation, vocabulary, grammatical structures and sentence construction, spelling, text binding, internal and external structure and adaptation, in students' own production and interaction.] (Lgy22, English 5, my translation)

The fact that neither the term *word knowledge* nor *vocabulary* is present in the content section *Reception* of the new curriculum is surprising since NAE dedicated an entire course section to vocabulary within the nation-wide *Läslyftet* (Reading enhancement), an initiative for continuing teacher education to en-

hance students' reading ability – that is their receptive skills – across all subjects (Skolverket, 2023). The course module called *Från vardagsspråk till ämnesspråk* (From everyday language to subject language) contained a section entitled *Ordförrådet – en framgångsfaktor* (Vocabulary – a success factor). More relevant for this thesis and even more surprising is the fact *academic vocabulary* is not mentioned in the curriculum, although it is given an important role in a chapter that teachers were required to read as part of this continuing education (Skolverket, 2017):

> Expansionen av ordförrådet under skoltiden sker i stor utsträckning i samband med undervisning i de olika skolämnena som alla tillför ordförrådet sitt speciella språkbruk med tillhörande facktermer. Ett viktigt tillskott sker också genom ett mer allmänt skolspråkligt ordförråd som spelar en allt viktigare roll för lärandet ju högre upp i årskurserna eleverna kommer. Det är ett ordförråd som kännetecknar språkbruket i all formell utbildning, liksom i mer formella sammanhang i samhället i stort. Det handlar ofta om abstrakta, ”akademiska” ord vars betydelse kan vara svåra att sluta sig till utifrån ordens form.
>
> [The vocabulary expansion during the years in school to a large extent takes place in connection with instruction in the various school subjects, all of which add to vocabulary, their special use of language and their associated disciplinary terms. An important addition is also the more general school-language vocabulary, which plays an increasingly important role in learning as students continue higher up in the school system. This is a vocabulary that characterises language use in all formal education, as well as in more formal contexts in society at large. These words are often abstract, *academic* words whose meaning can be difficult to infer based on the form of the words.] (Skolverket, 2017, p. 3, my translation)

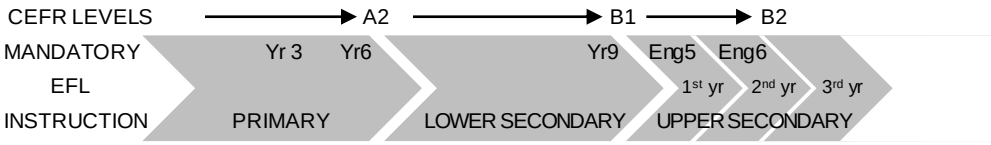
Considering the important role vocabulary in general and academic vocabulary seem to play for language development and educational communication, it is somewhat curious that vocabulary is not set out as a strand of reading instruction in English or as part of any grading criteria. On the other hand, considering this absence of vocabulary in the syllabi, it is unsurprising that English instruction and English textbooks in the Swedish school system have been found to be unrelated to vocabulary learning principles gained from vocabulary research; instead, the choice of what words to be included in instruction seems to be based on materials developers' and teachers' own intuition (Bergström et al., 2022a, 2022b, 2023; Nordlund & Norberg, 2020).

In connection with English grades, a pass grade in English from school year 9 is considered equivalent to the B1-level of the CEFR according to the National Agency for Education (Skolverket, 2021b). A pass grade in the final

mandatory upper secondary English 6 course is considered equivalent to the B2-level of the CEFR (Skolverket, 2015) (Figure 2).

Figure 2

Swedish syllabi for mandatory English instruction linked to the CEFR levels



Note. Inspired by Skolverket 2021a and adapted from Warnby (2023).

In the university-preparatory programmes at upper secondary education, the two mandatory courses, English 5 and English 6, share the same overall aim. For each course, the syllabus differs with regard to (*a*) the core content of instruction (Swe. *centralt innehåll*) and (*b*) the grading criteria (Swe. *kunskapskrav*). The core content is divided into three sections: (*i*) *Content of communication* (Swe. *kommunikationens innehåll*), (*ii*) *Reception* and (*iii*) *Production and interaction* (see Figure 1 in section 2.2.1.). This thesis focuses on academic vocabulary as a facet of academic English which is important to know for written receptive purposes. For this reason, the section *Reception* in the syllabi is especially interesting.

In the core content in the syllabus of English 5, there is no explicitly mentioned instructional guidance as regards the concepts of *academic English* or *academic vocabulary* (despite its central focus in the *Reading enhancement* course quote above). The section *Reception* includes content areas such as spoken English in varying contexts with variation in sociolects and dialects from different media, such as: fiction, songs and poems. Also included are strategies to perceive details and draw conclusions, searching for sources and source criticism, amongst other things. In one of the content areas for reception, *simpler popular science texts* are mentioned as one of many examples of texts that may be narrative, explanatory, discussing, argumentative and/or reporting:

> Talad engelska och texter som är berättande, förklarande, diskuterande, argumenterande och rapporterande – varje slag för sig eller i olika kombinationer. Till exempel intervjuer, reportage, manualer och enklare populärvetenskapliga texter.
>
> [Spoken English and texts that are narrative, explanatory, discussing, argumentative and reporting – each kind on its own or in combinations.

> For example, interviews, news coverage, manuals or simpler popular science texts.] (my translation)[5]

The core content in the syllabus of English 6 is similar to English 5 in that no guidance is explicitly given regarding *academic English* or *academic vocabulary*. Under the section *Reception*, several different content areas are mentioned, for example, spoken English in a relatively fast tempo, fiction including poetry and drama as well as contemporary work and excerpts from older works, strategies to draw conclusions, how attitudes and style are expressed in spoken and written English, etc. The word *simpler* in popular science texts in English 5 is omitted in English 6:

> Talad engelska och texter som är berättande, diskuterande, argumenterande, rapporterande och redogörande – varje slag för sig eller i olika kombinationer. Till exempel föredrag, debatter, formella brev och populärvetenskapliga texter.
>
> [Spoken English and texts that are narrative, discussing, argumentative, reporting and expository – each kind on its own or in combinations. For example, lectures, debates, formal letters and popular science texts.] (my translation)[6]

Returning to the notion of constructive alignment (see 2.2.1.), we have seen that the curriculum sets the objective for schools to prepare their students for university studies but that the English syllabi are far from clearly aligned with this principal objective; for example, there is no explicit mentioning of academic language features to be taught. A possible consequence of this lack of constructive alignment within the national curriculum documents is that teachers and schools will handle this unclear objective in vastly different ways resulting in large variations in students' proficiency in English for academic purposes. Additionally, based on such a consequence, one could hypothesise

[5] In the new curriculum, Lgy22, almost the same core content is expressed: "Talat språk … och texter som är instruerande, berättande, sammanfattande, förklarande, diskuterande, rapporterande och argumenterande,) …) Texter av olika slag och med olika syften, till exempel manualer, populärvetenskapliga texter och reportage" (Skolverket, 2022, p. 7). This means that the new syllabus adds *instructive* and *summarising* text types to the five text types in the former syllabus; furthermore, *interviews* from the older syllabus has been moved to belong to *spoken* instead of *written* language.

[6] In the new curriculum, Lgy22, there are slight moderations to the former curriculum: "Talat språk … och texter, även komplexa och formella, som är berättande, diskuterande, argumenterande, rapporterande och redogörande, … Texter av olika slag och med olika syften, till exempel formella brev, populärvetenskapliga texter och recensioner (Skolverket, 2022, p. 15). This means that the new syllabi add *complex* and *formal* to the five types of oral and written input; furthermore, *lectures* and *debates* have been removed and only include examples of oral input, not written.

that, although for some students the outcome in English 6 may be the same pass grade, this outcome may mean unequal preparation for academic English.

Related to the grades, a pass grade from English 6 – or an equivalent – is one of the requirements for admission to a Swedish university. Since English 6 is stated to be equivalent to CEFR-B2, upper secondary school students having passed English 6 may be assumed to have reached the level of English proficiency required for admission to many universities where the B2 level often is used (e.g., France, Germany, Norway, the UK).

2.2.4. English at university

English has become the *lingua franca* of academic communication (Mauranen, 2012; Mauranen et al., 2016) and the dual use of Swedish and English in Swedish higher education is common (Bolton & Kuteeva, 2012; Malmström & Pecorari, 2022; Nordic Council of Ministers, 2007; Young Academy of Sweden, 2022). During the past decade, the number of university courses and programmes with English as the medium of instruction (EMI) has increased (Airey, 2009; Malmström & Pecorari, 2022; Salö, 2010, 2016). However, in higher education in Sweden, some sort of parallel use of English and the majority language (here Swedish) is most common. This is especially true for English as a reading language even if the instruction is otherwise mainly in Swedish (Kuteeva, 2014; Malmström & Pecorari, 2022). This prominent role, presence and use of English in Swedish universities is also a relevant reflection of the position held by English at many universities globally (Bolton & Kuteeva, 2012; Dearden, 2015; Macaro et al., 2018; Mauranen et al., 2016; Montgomery, 2013)[7].

A central argument for the use of English in academic expanding circle settings is increased internalisation. Four explanations given for increased internationalisation of higher education are: academic, socio-cultural, political and economic (e.g., Malmström & Pecorari, 2022). For example, one major academic motive given is to increase the quality of operations, while an oft-quoted socio-cultural motive is to facilitate increased mobility. However, these motives often overlap, for example, the academic motive to recruit internationally also relates to the socio-cultural motive of mobility and intercultural understanding; the academic motive of spreading new knowledge internationally also entails a political motive in which the country aims to reinforce its position globally as a strong knowledge contributor; the economic motive of strengthening the country's economy depends on the academic motive of

[7] Also, beyond the scope of this thesis, certain upper secondary schools use English as medium of instruction for some or all subjects. The effects such immersion or content-language integrated learning programmes may have on, for example, students L1 Swedish academic literacy have been investigated (e.g., Lim Falk & Holmberg, 2016; Ohlsson, 2021).

producing high-quality and internationally attractive innovations (see Malmström & Pecorari, 2022, for a recent and detailed review of internationalisation in Swedish higher education). Moreover, all these types of reasons have direct implications for language use (Malmström & Pecorari, 2022). In 2018, an inquiry by the Swedish government reported on proposed objectives for increasing the internationalisation of Swedish higher education institutions (SOU 2018:3). Amongst many others, one proposition was to clarify the use of English and one of the consequences of internationalisation put forward was that higher education institutions may demand higher levels of FL proficiency for students and staff. Regarding English proficiency, the report concludes and assumes that:

> I dag är godkänt resultat på kursen engelska 6 (den näst högsta nivån i gymnasieskolan) obligatoriskt för grundläggande behörighet till högskolan, vilket innebär att alla som antas har en god kunskap i engelska.
>
> [Today a pass grade from the course English 6 (the next highest level in upper secondary school) is required to meet the general entry requirements to higher education, which means that all who are admitted have good knowledge in English.] (SOU 2018:3, p. 354, my translation)

It is evident that these internationalisation aims require a *lingua franca* for students, teachers and researchers to be able to communicate efficiently with each other. To this end, English – as the dominant language of academia – seems to be the most efficient (and unavoidable) choice both as a reading and publication language. However, as a reading language, English is merely the instrument for communicating disciplinary content (e.g., textbooks, articles) and little, if any, attention is given to language learning (Pecorari & Malmström, 2018). Although the use of English is very widespread and constantly increasing, differences do exist.

Two decades ago, a government report (SOU 2002:27) noted that the use of English within higher education varied, and was mostly used in natural sciences, technology and medicine. Furthermore, it was concluded that:

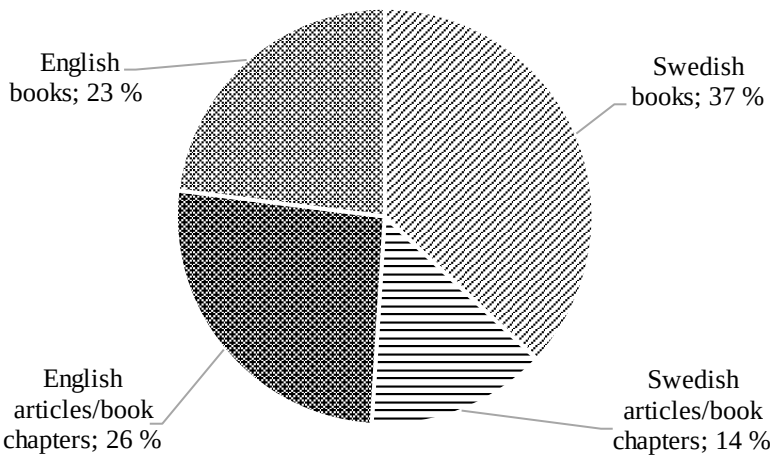
> Inom många ämnen är en stor del av kurslitteraturen engelskspråkig redan på grundutbildningsnivå, men framför allt ökar andelen på engelska drastiskt efter den första grundkursen.
>
> [In many subject areas, a large proportion of the assigned reading is in English already at the undergraduate level but increases drastically after the first undergraduate course.] (p. 80, my translation)

Twenty years later, a recent survey of 1706 reading lists across Swedish-medium of instruction courses in different academic disciplines/subject areas at

ten higher education institutions (five universities and five university colleges) showed that on average 65 percent of the undergraduate courses taught in Swedish have at least one obligatory assigned text in English (Malmström & Pecorari, 2022). Within the Humanities subject area, 84 percent of the undergraduate courses have at least one obligatory English text whereas the corresponding figure for Technology is 46 percent (Malmström & Pecorari, 2022) which is a change as compared to SOU 2002:27 (see above). Furthermore, in as much as 24 percent of these undergraduate courses with Swedish-medium of instruction, all the assigned course reading materials are in English. Irrespective of courses and disciplines, 49 percent of all the texts in the Swedish-medium courses are in English (Figure 3).

Figure 3

Share of all obligatory course literature in Swedish and in English respectively in Swedish-medium courses



Note. Figure adapted from Malmström & Pecorari (2022, p. 31).

Additionally, English is steadily reinforcing its dominance over the Swedish language as the publication language for research from Sweden. Twenty years ago, about 15 percent of doctoral dissertations were written in Swedish. Today that figure has decreased to around 7 percent. There is an inverted pattern for dissertations written in English with 85 percent in 2000, and 93 percent in 2019 (Malmström & Pecorari, 2022). The same relationship exists for published scientific articles, where, in 2019, about 5 percent were written in Swedish and 93 percent in English (Malmström & Pecorari, 2022).

The key message in this section is that a certain level of English academic proficiency is required (taken for granted) for almost any academic practitioner at Swedish universities and university colleges. Especially pertinent for an undergraduate beginner-student is an already established foundation of basic skills for academic reading in English.

2.3. Chapter summary

English has a strong position in Sweden across all societal levels and the distinction between English as an FL or an L2 is perhaps not meaningful. For this reason, the two terms will be used interchangeably. The English syllabi are CLT-oriented with little attention to vocabulary instruction. A pass grade from the final mandatory English course at upper secondary school, English 6, is demanded of students in university-preparatory programmes to be eligible for higher education studies. A central objective in the curriculum for upper secondary education is that all students develop adequate knowledge to be well prepared for tertiary level studies. Moreover, this chapter has confirmed that English is widely used at Swedish universities and university colleges. This is especially true for English as a reading language already at undergraduate level, showcasing the need to develop, as a preparation for taking on university tasks, an adequate academic English reading proficiency of which vocabulary knowledge is a central facet. The following chapter will focus on theoretical concepts regarding vocabulary research.

3. Vocabulary: Theoretical concepts

In order to frame the thesis and before describing the thesis' construct *academic vocabulary* in more depth, some theoretical concepts surrounding vocabulary knowledge, assessment and development will be outlined. This is a challenging task since "an overall theory of L2 vocabulary learning and knowledge is still missing" according to Bardel (2016) who argues that this is probably "a consequence of the complexity of the whole research area and the multi-faceted character of word knowledge" (p. 74).

A first matter of complexity is how to define words. To this end, the first section, 3.1., deals with word definitions and categorisations.

Another complexity is how knowledge of words can be conceptualised and assessed. Theoretical concepts within vocabulary knowledge and assessment will be clarified in the second section, 3.2.

Furthermore, vocabulary knowledge has been addressed as closely related to reading. Therefore, central matters related to vocabulary and reading will be presented in the third section, 3.3.

Moreover, notions related to vocabulary learning will be handled in the fourth section, 3.4.

Finally, this thesis positions itself as a contribution to knowledge of outcomes relating to the curriculum objective of university preparation. For this reason, the last section, 3.5., brings up some critical notions within the current communicative language curriculum and the role of vocabulary in curriculum planning.

3.1. Word definitions and categorisations

This section first presents how a *word* can be conceptualised. Thereafter, the notion of frequency will be described and how it can be used to group words together.

3.1.1. What counts as a *word*?

Within vocabulary research, a central aspect is what counts as a word since the *word* concept is vague and "has never been very clear in linguistic theory, although many different definitions have been formulated" (Bogaards, 2000,

p. 491). For a theoretical and linguistic review of the many aspects of a *word* see, for instance, Singleton (2000) or Wray (2015).

Related to the Saussurean terms *le signifiant* (that which signifies) and *le signifié* (that which is signified), an inevitable word aspect is that of the graphical/phonological *form* of a word and the ´true´ or conceptualised *meaning* of a word. For example, the graphical word *dog* has a 'true' meaning, that is, a meaning denoting a specific animal species. However, a word or a term can have several connotations and, therefore, be associated and conceptualised differently by individuals, for example, a *dog* can be conceptualised as 'a cute domesticated animal', as 'an aggressive nuisance', as 'an epileptic seizure anticipator' or as 'meat on the plate' (e.g., Hoffmann et al., 2018; Serpell, 2009). A solution to the dilemma with subjective associations to a word form described above is to say that the meaning of a word is that expressed in a dictionary. However, even in a dictionary a form may have several meanings, due to polysemy and homonomy; for example, the academic noun *study* has many meanings (Skoufaki & Petrić, 2021), for instance, 'the activity of studying', 'a piece of research', 'a drawing' or 'a room' (Collins COBUILD https://www.collinsdictionary.com/dictionary/english/study). Should we count these four meanings of the single form *study* as four words or do we treat the form *study* as one word having several meanings, some of which are more or less closely related? Furthermore, some words are more abstract than *dog*, for example, *reflect*, making clear-cut definitions more difficult. Moreover, some words carry very little meaning and are seldomly seen alone, for example, *a/an* + noun, yet one could easily think of *a/an* as one word.

Clearly, the definition of the word concept is tricky but important since it has implications for the methodology of categorising words, constructing word lists and assessing word knowledge. For instance, one may focus on single words (*enable, proportion, subsequently*) or multi-word units (*draw a conclusion, by and large, strongly opposed*). This thesis directs attention to single academic words. Additionally, single words can be counted as tokens, types, lemmas or word families.

A text consists of a number of *tokens*, that is, all the separate orthographic items, which also may be referred to as the running words of a text. However, every token is not semantically unique as the same token may often be repeated (especially high-frequency words such as *the, be, and, of* and *a*). Words that are graphically repeated are called *types*. Another way is to categorise words into *lemmas*, that is: "words with a common stem, related by inflection only, and coming from the same part of speech" (Gardner & Davies, 2014, p. 308). For example, the two graphically close tokens *word* and *words* can be grouped together under the noun lemma *word*, and the graphically distant tokens *is* and *were* can be grouped within the verb lemma *be*. Counting lemmas instead of tokens or types may be seen as reducing the lexical diversity in a

text which means that the 'words' a reader must know are fewer[8]. Sometimes a word form of a specific part of speech is identical to the word form of another part of speech, for example, a noun-lemma can have the same form as a verb-lemma (e.g., *study* and *rise*). A grouping of such lemmas is called a flemma (e.g., McLean, 2018). A fourth way of grouping tokens is by using the notion of *word family* (Bauer & Nation, 1993). A word family is "defined as a stem plus all closely related affixed forms" (Coxhead, 2000, p. 218), that is: a head-word and all its inflectional and derivational forms. For example, tokens such as *learner* and *learning* would be members of the word family *learn*, and *exposure* and *exposed* could be seen as members of the word family *expose*. The adoption of word family as the word counting unit decreases the number of 'words' in a text to a much greater extent than lemmas. The word family concept, however, has been contested when it comes to vocabulary assessment (See 3.2.3.).

Moreover, on the topic of single words, one could argue that, for example, chunked tokens such as *a lot of* or *word knowledge*, are representing one single lexical unit used as fixed multi-word units, for example: a formulaic sequence or a collocation. With reference to lemmas as well as word families, this thesis uses instruments targeting single academic word units shown to be typically frequent in academic discourse. The following section explains the principles of word frequency.

3.1.2. Frequency

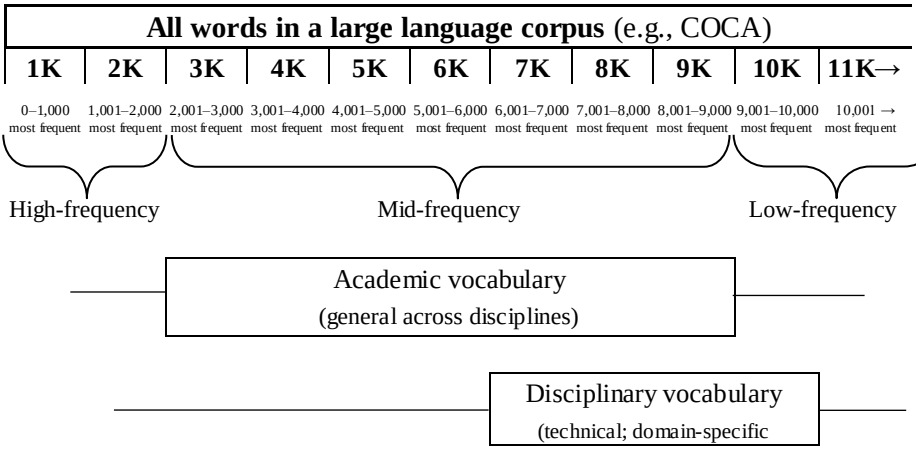
In vocabulary research, there are different ways to cluster words or define categories of vocabulary. A very common approach is to use a word's frequency in a relevant corpus as a basis, for example, the *Corpus of Contemporary American English* (COCA) as a comprehensive and representative body of English. By counting the number of occurrences of words within such a corpus, it is possible to define the 1,000 (1K) most frequently written words in that language followed by the next 1,000 words (2K), and so on (3K, 4K, etc.). From such frequency-based divisions, words can be clustered/classified as *high-frequency*, *mid-frequency* and *low-frequency* words (Masrai, 2019; Schmitt & Schmitt, 2014). High-frequency words (e.g., 1K–2K (Nation, 2013) or 1K–3K (Schmitt & Schmitt, 2014)) are the most common and general words occurring with a high frequency in all kinds of discourse. Conversely, low-frequency words occur more rarely at the 9,000 level and above (Schmitt & Schmitt, 2014). The lower the frequency of a word, the more likely it is that the word has a discipline-specific, more technical or specialised meaning. Depending on where the upper boundary for high-frequency vocabulary is set, *mid-frequency vocabulary* is placed somewhere from 3K/4K up to 9K (see,

[8] The inflected forms for a lemma are plural, third person singular present tense, past tense, past participle, gerund (*-ing*), comparative, superlative and possessive forms.

e.g., Nation, 2013; Schmitt & Schmitt, 2014) and has been shown to correlate more with reading comprehension than have high- and low-frequency words (Masrai, 2019). The construct under investigation in this study is *academic vocabulary* (also referred to as vocabulary of *English for general academic purposes* (EGAP) or *sub-technical* vocabulary) (Coxhead, 2016, p. 177). As will be set out in more detail in section 4.1., *academic vocabulary* is often contrasted with *disciplinary vocabulary* (also referred to as vocabulary of *English for specific academic purposes* (ESAP) or *technical* vocabulary). *Academic vocabulary* consists of words that often occur across a wide range of disciplines without being too general or too technical and it encompasses a broad range of words with respect to frequency. As such, academic vocabulary can be seen as centered alongside the mid-frequency register, however, it cannot be labelled mid-frequency since there are mid-frequency words that are not typical for academic discourse. Figure 4 is intended to illustrate the above-mentioned definitions.

Figure 4

Illustration of vocabulary frequency bands, frequency registers and academic and disciplinary vocabulary



Having addressed definitions of a *word* and vocabulary frequency categorisations, the following sections deal with aspects and issues regarding word knowledge and vocabulary assessment.

3.2. Vocabulary knowledge and its assessment

In this section, some theoretical concepts regarding word knowledge will be presented in the first sub-section. As vocabulary assessment is inevitably linked to the concepts of what kind of word knowledge is targeted, some specific assessment concepts will be presented in the second sub-section.

3.2.1. Word knowledge

How words are known, stored and retrieved in the mental lexicon, that is, in humans' internal repository of words, is yet to be understood. As with any kind of internal knowledge and comprehension, we need some observable act of comprehension to enable us to draw inferences about cognition, or as stated by Robinson (1988) "the preoccupation with a 'lexicon' itself … ; with the modular organisation of 'words in the mind' … , obscures the fact that we have no access to such a repository except via 'words in the air' or 'in the text'" (p. 2). For this reason, the term and concept *mental lexicon* has been criticised for having little to offer research (e.g., Elman, 2011). Despite such criticism, the term *lexicon* highlights the existence of a difference between the mental representation of word meaning (which relates to declarative knowledge; what we know) and the observable acts of uses and negotiations when confronted with word form – spoken or written (which relates to procedural knowledge, what we can do) (for a discussion of declarative and procedural vocabulary knowledge, see, for example, Robinson, 1988, 1993[9]).

From the above, it is obvious that attempts to describe word knowledge entail descriptions of what a learner can do with words, for example, Cronbach (1942) summarised, into five aspects, "(v)arious sorts of behavior that may be called for in understanding a word" (p. 206): generalisation (having to do with defining words), application (having to do with connecting forms and meanings), breadth of meaning (distinguishing different meanings of a word), the precision of meaning (discriminating between nuances of different words) and availability (making "use of the concept in … thinking and discourse" (p. 207).

Related to availability is fluency which is an aspect of word knowledge, which, for reception, means that the word's meaning can quickly and automatically be retrieved from the mental lexicon. This kind of vocabulary knowledge is sometimes called sight vocabulary (Coady & Nation, 1988; Laufer & Aviad-Levitzky, 2017). In reading, for example, fluency has a significant influence on the time it takes for an L2 learner to identify and decode a word phonologically/orthographically and morphologically, to retrieve the

[9] Out of topic for this thesis' investigation but with regard to vocabulary knowledge and use, the interested reader will note that Robinson (1988) also makes a distinction between *knowledge* and *skill* relating this to the difference between, for example, *competence* and *performance.*

word in the mental lexicon, to recall its meaning – perhaps through an L1 translation – and then, to act upon this process according to a task. Automatised word knowledge is thus important for freeing working memory for cognitively more demanding processes (see below section 3.3).

Laufer (1991) outlined five components of word knowledge which all have their typical pitfalls for a learner: (*i*) form, (*ii*) word structure, (*iii*) syntactic behaviour, (*iv*) meaning and (*v*) lexical relations. In the *meaning* component, Laufer includes not only the a) referential meaning but also b) the affective meaning of a word, c) the pragmatic meaning of a word and d) the frequency of a word. The *lexical relations*-component considers the meaning relationships between words where synonyms and hyponyms are examples of such lexical relations. With regard to Laufer's terminology, the tests used in this thesis can be said to tap into the referential meaning of word knowledge and the lexical relations-component.

Also in an attempt to conceptualise word knowledge, Nation (2001, 2013) presented a table (Table 1) that has become influential for our understanding of what is involved in knowing a word.

Table 1

What is involved in knowing a word

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

Note: R = receptive knowledge, P = productive knowledge. From Nation (2013, p. 49).

Nation divides vocabulary knowledge into three levels (*form*, *meaning* and *use*), and nine aspects (e.g., *spoken, form and meaning,* and *grammatical functions*). Table 1 can thus be seen as a taxonomy hierarchically ordering different levels of word knowledge, that is: more advanced levels of word knowledge occur lower in the table. Furthermore, each aspect has its own particularities separated whether they concern receptive or productive knowledge, where the productive skills are presented after the receptive indicating the more demanding characteristics of production in comparison to reception. A basic aspect of word knowledge is the simple connection between a word's form (spoken or written) and its meaning. However, even this basic

aspect becomes problematic when considering the several meanings certain word forms have, for example, polysemy.

Knowing several meanings of one word form may sometimes be referred to as depth (having deeper knowledge of a word form), whereas how many words that are known often is referred to as breadth/size but this distinction is not clear-cut and has been criticised (Gyllstad, 2013; Read, 2004). In Nation's table above, the 'spoken' and 'written' word knowledge aspects of the 'Form'-level and the 'form and meaning'-aspect of the 'Meaning'-level are often viewed as aspects of breadth, "whereas the remaining ones … are usually considered depth aspects" (Gyllstad, 2013, p. 18). However, a collocation such as *fast food* and *hard work* can be defined as a single-word unit, which is part of breadth according to Nation's table. Also, a typical measure of word meaning (i.e., breadth) is to link a word's meaning with another word, and such a linking can be viewed as associations in Nation's table (i.e., normally seen as depth). A Nationesque view would be that a word has one core meaning from which a learner builds his/her understanding, which is a definition that can be related back to Cronbach (1942) who talked about a learner's total vocabulary as measured by the knowledge of the "commonest meaning of each word" (p. 209).

The depth dimension may be more crucial for production than for reception. For receptive purposes, it may be sufficient to have knowledge of the core or commonest meaning to make meaning out of a text. This is where partial knowledge comes in, which refers to the fact that a learner may only have partial knowledge of a word's full array of meanings or uses. For an L2 learner, it is a daunting challenge to master all possible and precise meanings and uses of a word. Even an L1 user would be likely to exhibit partial knowledge of a number of academic words. This may be especially true for an L1 user who is inexperienced in academic discourse (Nation, 2001, 2013).

Since this thesis uses academic vocabulary as an important facet of academic reading proficiency, it is the receptive aspect of the form-meaning connection in Nation's table (Table 1), that is in focus. However, receptive knowledge should also be separated into oral or written communication. Written discourse differs from oral in many aspects, for example, oral communication uses more high-frequency vocabulary than written. For this reason, it is important to clarify that this thesis assesses knowledge of words that are typically encountered when reading written academic communication, hence the notion of written receptive academic vocabulary knowledge.

3.2.2. Assessing vocabulary knowledge

Fundamental issues within measurement theory apply to vocabulary assessment, for example, defining the construct, arguing for the reliable and valid

uses of the measurement instrument, as well as acknowledging threats to validity. This section will briefly address some of those theoretical notions on a general level as they relate to vocabulary assessment.

Validity is probably the most central aspect of assessment (Crooks et al., 1996). A simple view of validity is that a test measures what is intended and nothing else. However, the complexity can be reflected by the numerous concepts of validity that exist, for example, concurrent validity, predictive validity, content validity, construct validity, to name a few (Kane, 2006; Lissitz, 2009; Messick, 1996; Milton, 2009; O'Sullivan & Weir, 2011; Xi & Sawaki, 2017). *Concurrent validity* refers to how performance results on similar constructs are correlated. A specific use could be when validating a newly developed test by correlating and comparing it with an earlier standard. *Predictive validity* concerns how adequate test results may be to predict a test taker's future performance[10]. Examples of performances that are used to predict possible future academic success are the Swedish Scholastic Aptitude Test (*Högskoleprovet*) or the International English Language Testing Service (IELTS). *Content validity* examines the extent to which a test consists of sufficient and relevant content to measure what is intended. A test of academic vocabulary should include an appropriate set of items that represent the targeted domain. *Construct validity* generally refers to a holistic view on validity and the interpretations and inferences about the construct that can be made based on test results. There are two major threats to construct validity: 1) *construct underrepresentation* happens when a test consists of too narrow a set of items and therefore does not capture the main dimensions or important aspects of the construct; 2) *construct-irrelevant variance* is when the test is too broad and, thus, measures additional constructs that are irrelevant for the purported construct/domain (Messick, 1996).

A theory on validity is the basis for the assumptions about a test, whereas a validation model enables the operationalisation of the validity theory giving meaning to test results (O'Sullivan & Weir, 2011; Xi & Sawaki, 2017). A theoretical challenge in such a validation pertains to the validity arguments which, in social science research, are hard to prove empirically (Crooks et al., 1996; Kane, 2006). Reliability measures are often used as a validity argument when, in fact, they are only one but an essential part of a whole validation process (Crooks et al., 1996; Messick, 1989) and, although seemingly objective, they are, in fact, value-laden and dependent on subjective choices (Gipps, 1999). Arguing for the valid use of a test and assessing threats to validity "require(s) conceptual analysis and professional judgement" (Crooks et al., 1996, p. 282). In the following paragraphs, some of the above-mentioned concepts will be related to vocabulary assessment specifically.

[10] Both *concurrent* and *predictive* validity are often considered forms of *criterion-referenced* testing as opposed to *norm-referenced* testing (e.g., Brown & Hudson, 2002).

First, when assessing vocabulary the construct has to be defined (e.g., Read & Chapelle, 2001). The central questions are: What kind of *vocabulary* is to be tested (section 3.1)? What kind of vocabulary *knowledge* is targeted (section 3.2.1)? The answers to those questions will always depend on the purpose and the context (Schmitt et al., 2020). For instance, if the purpose is to give an indication of a learner's knowledge of words relevant for reading, some type of written receptive knowledge may be adequate (i.e., neither spoken nor productive). If the context includes beginner L2 learners, the words could be collected from a corpus of elementary level graded English written texts (i.e., not advanced) and target words up to 2K. In this case an assessment tool could be the 2K level of the *Vocabulary Size Test* (VST) by Nation and Beglar (2007) or the 2K level of the *Vocabulary Levels Test* (VLT) originally developed by Nation (1983), revised by Schmitt et al. (2001) and updated by Webb et al. (2017). Both these tests are widely used, they are built on the word family concept, and they have a multiple-choice format to test meaning recognition (see item examples in Figure 5 and Figure 6).

Figure 5

Example of a VST 2K item (Nation & Beglar, 2007)

STONE: He sat on a **stone**

a. hard thing
b. kind of chair
c. soft thing on the floor
d. part of a tree

Note. Item example from the VST by Nation and Beglar (2007).

Figure 6

Example of a VLT 2K item

1 attack 2 charm	_____gold and silver
3 lack 4 pen	_____pleasing quality
5 shadow 6 treasure	_____not having something

Note. Item example from the VLT by Schmitt et al. (2001).

Conversely, for advanced learners or students aiming at university studies, academic vocabulary may be more appropriate to measure than 2K. Which vocabulary construct to assess must be argued for, and there is no distinct measure to rely on. Therefore, in an argument-based approach, one has to offer arguments that address the definition of the construct and demonstrate that the chosen test measures this delineated single construct in a unidimensional way.

The issue of unidimensionality has, however, been debated when it comes to vocabulary testing (e.g., Schmitt et al., 2020). Knowing the meaning of one tested word does not necessarily imply knowledge of another discrete word item from the same frequency band, for example, *clever* is followed by *interactive* in the 6K band (Schmitt et al., 2020, p. 115). Nonetheless, vocabulary assessment often relies on some sort of vocabulary categorisation for unidimensionality, for example, through frequency (e.g., 1K, 2K, etc (e.g., Nation & Beglar, 2007)), or through the use of thematic clustering of words (e.g., *frog, pond, hop, swim, green* as parts of the thematic schema 'frog' (e.g., Tinkham, 1997)); or according to a specific vocabulary typical for a genre, for example, academic vocabulary (e.g., Pecorari et al., 2019). Factor analysis is often used in measurement theory to check whether the test is unidimensional or contains additional assessed domains. Other statistical analyses can also be applied to the data to see whether test items may affect reliability or be biased. However, such statistical measures may only help to detect threats to validity; ultimately, the choice of what kind of vocabulary to measure depends on subjective argumentation. This also applies to the choice of what kind of vocabulary knowledge is to be tested.

A basic aspect of written receptive word knowledge is, as described in 3.2.1., the linkage of a word's form with its meaning. However, it is hard to observe and measure this internal linking process in the mental lexicon, that is: the measurement of the process by which a person receptively encounters a word, then mentally and cognitively connects meaning to the word form demands, what can be called, an external act of comprehension. Thus, to measure receptive knowledge, some sort of action is required indicating the person's comprehension (Milton, 2009). In the case of the VST and the VLT (see item examples above), this act of comprehension is produced by combining the target word's form with a given definition or a synonym (the synonym connection would correspond to the aspect of association in Nation's table above, 3.2.1.). However, in a meaning-recognition test we can only observe the responses on the test items, and, depending on the assumption of how words are stored mentally, other measurement issues may arise.

As laid out in 3.1.1. above, words are often categorised into lemmas or word families. A theoretical assumption about word knowledge is that we chunk words together in some way to store them in our mental lexicon (Levelt, 1989). The lemma and the word family concept build on such an assumption. Word families are frequently used as the counting unit in vocabulary tests, for

example, the VLT. A learner who knows one word in a word family is assumed to also know, at least receptively, the meaning of other word family members or infer their meaning from the known word family member (Bauer & Nation, 1993). This may be true for "low-level learners who read simple texts" (Laufer, 2021, p. 967). The word family notion is questionable since there may be large differences between two word family members, for example, in frequency or in morphology. In the Academic Word List (AWL; Coxhead, 2000), a headword like the verb *interpret* is much more frequent and morphologically less complex than its word family noun sibling *reinterpretation*. Derivational and morphological knowledge of a headword may be considered a different construct than meaning-recognition knowledge of the headword (Leontjev et al., 2022). Furthermore, a word family often contains words from different parts of speech, as in the example in the previous sentence.

The theoretical reasoning behind the use of word families as the vocabulary testing unit is similar to that of lemmas as regards the extrapolation of knowledge outside the tested word. When using lemmas as the word-unit it is assumed that a test taker who knows one lemma also knows other forms of that lemma, for example, all verb forms, including: *interpret, interprets, interpreted* (Milton, 2009). Differences between a lemma's different forms are not as disparate as they can be between different word family members. It may be logically easier to assume that a test taker who exhibits knowledge of *interpret* may also understand *interpreted* in context, but not necessarily *reinterpretation*.

These theoretical and methodological problems with assuming knowledge of other words than the ones observed in a test and with chunking words together from different frequency bands or from different parts of speech "should be possible to resolve first of all by counting lemmas instead of word families" (Bardel, 2016, p. 100). In fact, to date, several scholars suggest that the use of /f/lemmas is a more adequate word-counting unit than the word family in vocabulary assessment (e.g., Kremmel, 2016; McLean, 2018; Stoeckel et al., 2020) and especially for receptive purposes (Brown et al., 2022). Nonetheless, it is worth noting that testing a lemma or a word family's headword entails testing single words and not other word definitions (for example collocations).

Both the VST and the VLT target single words, however, the VST places the target word in a neutral example sentence, whereas the target words in the VLT have no context. Assessing single words out of context may seem nonauthentic if test results are to be used as predictors of larger language constructs, such as reading performance (e.g., Read, 2000; van Zeeland, 2013). However, as soon as context is entered into a vocabulary test, additional errors of measurement may be introduced, for example, guessing from contextual clues. The VLT test format gives no context to the target words but suffers from other potential threats. In the VLT test format, a definition/synonym is to be matched with the best corresponding single-word alternative (from a list

including both the target word and distractors) (see Figure 6 above). Such a test format singles out contextual random effects outside the *form – meaning* combination, but since three definitions/synonyms should be matched with three words item exclusion strategy becomes a potential issue. Inevitably, there are always errors of measurement. Two important threats to validity in tests like the VLT will be mentioned, sample size and guessing.

First, defining a construct in language assessment is challenging and often includes a complex construct (Milton, 2009). For vocabulary assessment, testing all words of a vocabulary register, for example, all one thousand words in the 3K band, "is not possible" (McLean, 2021, p. 127). It is therefore important to make a sample of items representative of the whole construct, that is: the content of the test must be valid for the purpose of the test. It is well known that if we increase the number of items, we also increase the internal consistency (Cronbach's alpha, for instance). Conversely, when a complex construct is represented by a small number of items, we risk having construct under-representation. Following this logic, a test with 30 items representing a total construct of 705 items (i.e., 4% of the construct is sampled in the test) may seem more robust than a test of 57 items representing a total construct of 1,908 items (i.e., 3% of the construct is sampled in the test)[11]. However, the internal consistency would normally be better with the longer test and reported as a more reliable test (see, e.g., Crooks et al., 1996).

The item sample size is thus an important aspect because a test taker can have bad luck or good luck on the testing day. Imagine two students knowing half of the words in a word list but not knowing the same words. The first student may have bad luck and score zero since the test only tested words from the half he/she did not know. The other student may have such good luck that all tested items are known and therefore get a full score. One scores zero, and the other one gets a full score, despite both students having only knowledge of half of the underlying word list. Normally, the theoretical logic behind these tests assumes that a learner scoring 15 on a 30-item vocabulary test measuring knowledge of a 1,000 word band is estimated to know 500 of the 1,000 words. This is a strong theoretical assumption given the issue of bad or good luck with the sampled items as a result of a small item sample size.

By making the impossible possible (see McLean, 2021, above), Gyllstad together with Stewart and the above-mentioned McLean investigated this assumption empirically (Gyllstad et al., 2020). They let 103 Japanese EFL university students take two different 1000-item tests measuring all the words in the 3K-band (how long it took students to answer all 2,000 items and whether it was administered in one session remains unknown; students were paid per

[11] The numbers 705 and 1,908 are calculations of word families when the Compleat Web Vocabulary Profiler v.2.6 (see section 4.2.) defines and counts the word families when entered with all the ≈3,000 items from the Academic Word List (Coxhead, 2000) and the ≈3,000 items from the Academic Vocabulary List (Gardner & Davies, 2014).

hour, had no time restrictions and answered through an online link). This remarkable dataset made it possible for the researchers to make bootstrap samples for each test by testing out different test lengths (10, 20, 30, 40, 60, 100 and 200 items) and comparing those bootstrap samples with the participants' observed scores on the full 1,000 item test. Increasing the number, of course, decreased the average differences, thus, the confidence intervals became smaller. For the meaning-recognition test format, increasing the number of items from 30 to 60 or 100 did not change the confidence interval slopes to any large degree. Therefore, Gyllstad et al. (2020) concluded that although "it is difficult to establish a very firm cutoff point in terms of the number of items, our data seemingly indicate that a sample size around 30 or more items per 1000-word frequency band provides a decent accuracy, and at the same time constitutes an item number that is practically conceivable" (p. 572), that is: a 3 percent sample size.

Item size is thus an important threat to validity and the examples presented above show the importance of considering the item sample size in testing situations, especially if scores have meaning on an individual level. However, increasing the sample size of participants may equal out some of these errors when estimating group mean scores (Gyllstad et al., 2020). Group mean scores become thus more reliable since they will include test takers with bad luck as well as those with good luck.

A second threat to validity concerns guessing in a meaning-recognition test such as the VST or the VLT. In a multiple-choice or matching test format, a guessing effect should always be considered. A correctly observed response on an item may represent true knowledge or some sort of guessing, thus introducing some level of construct-irrelevant variance. Different procedures to estimate the size of this guessing effect have been suggested. Using Item Response Theory with large sample sizes, a three-parameter logistic (3PL) model may estimate each test item's guessing parameter (Embretson & Reise, 2013; Hambleton & Swaminathan, 2013).

Another procedure is to use a combination of a meaning-recognition test and a meaning-recall test (e.g., Gyllstad et al., 2015; Pecorari et al., 2019; Schmitt et al., 2001). By comparing test takers' meaning-recognition responses on a word-item with their meaning-recall responses on the same word-item, a guessing effect was considered. This procedure is based on the theoretical assumption that if the test taker cannot recall the meaning of the word but has responded correctly on the meaning-recognition format, the meaning-recognition response is due to guessing. For this reason, the scores on tests like the VLT are said to be inflated by some level of guessing, that is, the true score is lower than the observed score (e.g., Gyllstad et al., 2015; Stewart & White, 2011). Stewart and White (2011) estimated an average VLT score inflation of almost 17 percent as a consequence of guessing.

Furthermore, proponents of the meaning-recall format advocate using meaning-recall instead of meaning-recognition as a better measure of reading

comprehension largely because meaning-recall measures in relation to meaning-recognition measures have been shown to correlate higher with reading test scores (McLean et al., 2020; Stewart et al., 2021; Zhang & Zhang, 2020). However, comparing responses on a receptive test with responses on a productive test changes the construct definition (Laufer & Aviad-Levitzky, 2017; Laufer & Goldstein, 2004; Leontjev et al., 2022; Read, 2000). A meaning-recall test demands productive word knowledge and shows how a learner externally produces the meaning of the word in oral or written creative tasks, either in their L1 or L2. This is in contrast with the meaning-recognition test which attempts to capture receptive word knowledge, that is without having to produce language. As described in section 3.1.2 the latter is at a lower level of word knowledge than the former. However, the idea of word knowledge being multidimensional (for example, meaning-recognition vs. meaning-recall) has recently been contested, and, instead, it is argued that L2 vocabulary knowledge is unidimensional irrespective of the tested word knowledge aspect (González-Fernández, 2022). Nonetheless, “the discrete description of word-knowledge aspects still has a value" (González-Fernández, 2022, p. 1148), and, as mentioned above, meaning-recall as well as meaning-recognition formats are often used as measures of the receptive task of reading comprehension.

3.3. Vocabulary and reading

This section will first lay out some basic assumptions about reading ability and the role vocabulary plays in reading comprehension thus providing validity arguments for using vocabulary measurements as indicators of reading comprehension. Thereafter, concepts and estimations of how many words a reader must know for text comprehension will be presented.

3.3.1. Basic assumptions about reading comprehension

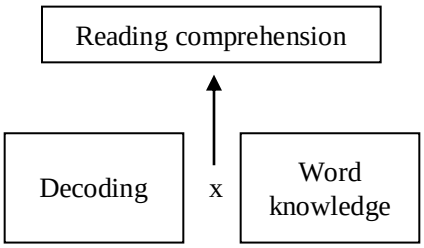
In the following paragraphs, three basic assumptions about reading comprehension are briefly presented with the aim of offering validity arguments for using measures of vocabulary knowledge as an essential component of reading comprehension.

First, a straightforward framework for conceptualising the process of understanding written language is the *Simple View of Reading* (Gough & Tunmer, 1986), according to which two major components are required for reading comprehension: word decoding and word knowledge (Figure 7). There is substantial agreement in reading research regarding the validity of the Simple View of Reading (Alderson et al., 2016; Hoover & Gough, 1990; Hulme &

Snowling, 2011; Perfetti et al., 2005; Stuart et al., 2008). As learners grow older, reading comprehension depends increasingly on the word knowledge component and to a lesser extent on the decoding component (Gough et al., 1996). For the average targeted participant in this thesis, the word decoding component (deciphering grapheme-phoneme couplings, i.e., the reading code) is not considered important as a measure of their reading comprehension.

Figure 7

Illustration of the Simple View of Reading

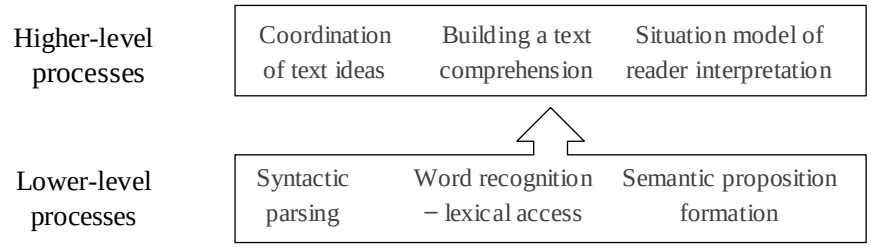


Note. The illustration is based on Gough and Tunmer (1986).

Second, a prominent distinction to be made about reading is between *lower-level* (identification) and *higher-level* (interpretation) reading processes (Grabe, 1991, 2009; Grabe & Stoller, 2019) (Figure 8). "The most fundamental requirement for fluent reading comprehension is rapid and automatic word recognition" - also called lexical access – which is included in the lower-level reading processes (Grabe & Stoller, 2019, p. 16). More demanding skills are included in the higher-level processes, for example, building a model of text comprehension in the reader's head; "as the reader continues to build an understanding of the text, the set of main ideas forms a mental representation of text comprehension" (Grabe & Stoller, 2019, p. 23). Since *word recognition* is included in the lower-level reading processes and, thus, is a requirement for building a *model of text comprehension*, the use of word meaning-recognition measurements in this thesis is valid as a minimum requirement for text comprehension to take place.

Figure 8

Lower-level and higher-level reading processes

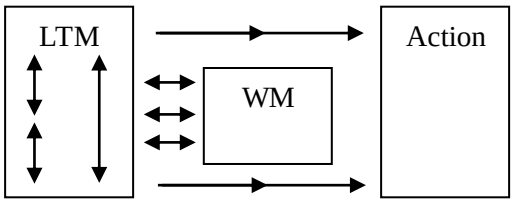


Note. Adapted from Grabe and Stoller (2019, p. 16).

Third, related to the above, the execution of tasks, for instance, understanding written texts, involves complex processes and multiple links between the working memory (WM) and the long-term memory (LTM) (Baddeley, 2012) (Figure 9). The LTM represents the individual's storage of memorised words – the mental lexicon – whereas WM represents the handling of this mental lexicon – sifting out the meaning from the process of putting words and their meaning together, that is, the reading process. When the WM goes to the memory store (LTM) to pick out relevant word meanings that fit into the context, it is an active and online process[12]. Thus, a valid argument for testing word knowledge is that limited word knowledge will occupy too much of the WM's online capacity, so text comprehension is unlikely to happen.

Figure 9

Illustration of the links between working-memory (WM), long-term memory (LTM) and action



Note. Adapted from Baddeley (2012, p. 18).

[12] The WM – or the *central executive* in Baddeley's words – can be seen as a "series of fluid systems that require only temporary activation" whereas the long-term memory represents "more permanent crystallized skills and knowledge" (Baddeley, 2012, p. 11). Crystallised word knowledge (in the LTM) is required for allocating place in the working memory to enable fluid processing of the text at a more abstract level.

Understanding a sentence requires the working memory to have access to words and their meaning and to keep this information momentarily until the linguistic relation between distal parts of the sentence can be established, thus, making comprehension on the sentence level possible. In the example sentence (1) from Lewis et al. (2006, p. 448), the relationship between the noun *toy*, in the subordinate clause, and its verb *arrived* is separated by a prepositional phrase belonging to the noun itself, namely *from her uncle in Bogotá*.

(1) Melissa knew that the toy from her uncle in Bogotá arrived today

The prepositional phrase entails a storage interval, which means that the head noun *toy* must be held in the working memory until the reader reaches the verb *arrived*. "The basic idea is that each incoming word triggers retrievals to integrate that word with the preceding structure" (Lewis et al., 2006, p. 449). The example sentence may seem easy to understand in terms of frequency, but if, for example, the word *arrived* is unknown, the reading process will be stopped, the word must be looked up in a dictionary, and the storage interval between this point and the preceding point of the sentence is likely to be too long, thus, the meaning and the context is lost. This can be described as a continuous and online processing of written content.

Example (2) below is used to illustrate academic language and is a sentence from Masrai and Milton (2018, p. 46). It represents a more complicated text than example (1) above: Example (2) includes more storage intervals and less frequent and more advanced vocabulary:

(2) However, an understanding of the word *demand*, even in its most general form, ought to allow commerce-related expressions, such as *demand curve*, to be at least partially understood.

Here, for example, a storage interval is required between the subject *an understanding* and the verb phrase *ought to allow*, and between the verb *understood* and its distal head noun *expressions*. Furthermore, linguistic relations must be established not only within the sentence but also between sentences and passages in the whole text, as indicated by the word *however*. Additionally, it may be interesting to know that the words *however, understanding, demand, general, form, related* and *partially* are academic words according to Gardner and Davies (2014). Knowledge of words is thus important for giving room to the working memory to establish relations between different parts of the text.

The above-mentioned concepts are valid for L1 as well as L2 reading. However, L2 reading may be even more demanding because it may require some sort of translation process between L1 and L2. First, as regards the Simple view of reading, decoding L2 letters that do not exist in the L1 may hinder reading comprehension. Second, any inability to recognise L2 words at the

lower-level reading processes will negatively affect higher-level processes. Third, looking up the meaning of unknown L2 words in a dictionary will occupy the working memory that has to set aside other temporarily stored information about, for example, the meaning of the overall text passage. Thus, such translation processes will occupy a substantial part of the WM processing capacity rather than dealing with higher-order understanding of the text.

The above-mentioned concepts all provide arguments for using measures of word knowledge as a minimum requirement for reading comprehension. In the following section, the issue of how many words a reader has to know to enable text comprehension is addressed.

3.3.2. How many words are needed to understand a text?

From the above account it is evident that L2 reading is a complex, variable and internal process. For this reason, the construct of reading ability is never fully reflected in observed scores of reading comprehension tests (Alderson, 2000; Alderson et al., 2014; Grabe, 1991). Reading ability includes several variables like metacognitive reading strategies[13], motivation for reading, content knowledge, and language knowledge of which a well-developed lexicon constitutes the lion's share. Although a learner's grammar knowledge also affects the reading process, vocabulary knowledge is more strongly associated with reading comprehension than grammatical knowledge (Aryadoust & Baghaei, 2016), or as Laufer and Aviad-Levitzky (2017) put it:

> It has … been shown that vocabulary knowledge is highly correlated with reading in one's first language … and even more so with reading in an L2 … . We acknowledge that vocabulary knowledge is not the only condition for successful comprehension. However, we are not aware of studies that show that L2 learners with poor (below threshold) vocabulary achieved good reading scores because of good grammatical knowledge. (p. 730, incl. footnote 1)

Vocabulary knowledge is thus the key predictor of L2 reading comprehension (Laufer, 1992; Laufer & Ravenhorst-Kalovski, 2010; Laufer & Sim, 1985; Leeming, 2014; Qian, 1999; Schmitt et al., 2017; Schmitt et al., 2011; Staehr, 2008) and is central for academic literacy (Qian, 2002) and academic success

[13] Regarding metacognitive reading strategies, this level of processing is likely to be out of reach for an individual who has insufficient levels of word knowledge, which, in turn, results in him/her remaining on the first level reading processing. For a learner with a limited vocabulary of barely the high-frequency words, Milton (2022) argues that it is a "mistaken advice that unknown words can be avoided through strategy use" (p. 167).

(Daller & Phelan, 2013). In short, L2 reading development "centrally involves learning words" (Grabe & Stoller, 1997, p. 119) since a reader needs to master many thousands of words and retrieve their meaning automatically and quickly before being able to, for example, guess unknown words from context. However, a large academic vocabulary is not enough to guarantee academic reading comprehension, but insufficient vocabulary size can cause severe reading problems. In fact, vocabulary is such a good predictor of the larger reading construct that Coady and Nation (1988) argue that although "there are a number of factors in a text which contribute to its ease or difficulty for a given reader, … we can most accurately predict that fact by measuring one variable, vocabulary, and extrapolating from it to the overall case" (p. 97–98). In short, using measures of word knowledge is a valid way of capturing the most central component of reading comprehension.

Since vocabulary is so important for reading comprehension, many scholars have tried to estimate how many words a reader must know to understand a certain type of text. As always with assessment, such estimations will depend on the purpose and the context. The purposes of reading will define what is demanded of our understanding of the text. Reading to understand a text on a global level may demand other vocabulary knowledge than reading to understand details; reading for pleasure may demand other vocabulary knowledge than reading for solving a problem. The context is important since different genres use somewhat different lexis. One would expect to read the formulaic sequence *once upon a time* when reading a fairy tale, whereas one would be very surprised to encounter it in an academic paper (unless as a quote, for example, in children's literature studies). Conversely, one would not expect to meet a word like *section* in a fairy tale but would not even reflect upon its occurrence in an academic text. Thus, to answer how many and which words a learner must know is a relevant question except for one aspect: the reader of a text does not have to know words that do not occur in the text. Fittingly, the vocabulary research field has defined some notions regarding this question.

The *lexical threshold*, which has been identified as the minimal vocabulary knowledge demanded for reading comprehension and before other metacognitive reading comprehension strategies can be activated (e.g., Hu & Nation, 2000; Laufer, 1989, 1992; Laufer & Ravenhorst-Kalovski, 2010; Nation, 2006; Schmitt et al., 2011). The lexical threshold is a theoretical notion dependent on two things: (*i*) lexical coverage and (*ii*) receptive vocabulary size.

Lexical coverage may refer to how many words in a text are known by a reader. If a reader has 95 percent lexical text coverage this indicates that the reader has lexical knowledge covering 95 percent of the running words in the text. Lexical coverage may also refer to the word registers of a text. For example, we can say that the 3,015 lemmas in the Academic Vocabulary List (AVL; Gardner & Davies, 2014) provide 10–14 percent lexical coverage of an academic text. This means that 10–14 percent of all tokens in the academic text are covered by the AVL list. Knowing all the AVL words means that the

reader knows at least 10–14 percent of the running words in the academic texts.

Receptive vocabulary is a notion that can be used to denote all those words for which a learner can retrieve the meaning effortlessly (sight vocabulary) or with the aid of contextual clues (comprehension vocabulary) (Laufer & Aviad-Levitzky, 2017). A receptive vocabulary size that is adequate to the text genre facilitates reading comprehension and reading fluency. The lexical threshold, thus, depends on the reader's vocabulary knowledge and the lexical demands of the text to be read. For a reader who barely knows a handful of academic words, the vocabulary load is likely to impede understanding; the lexical threshold will simply be too high for this reader to understand this type of text. Estimates of how big a receptive vocabulary is needed for reading and understanding different types of language have been a central interest of previous studies.

It is generally accepted that a reader needs to know 95–98 percent of the words in a text in order to understand the content (e.g., Hu & Nation, 2000; Laufer & Aviad-Levitzky, 2017; Nation, 2001, 2013). Not knowing 5 percent of the running words in a text means not knowing one word in every 20 words, which is, on average, one word every second line, and only some learners will be successful in reading and understanding a text with this lexical coverage, that is, 95 percent lexical coverage. An increase to 10 percent of unknown words will stop the reader on every line and clearly hinder reading comprehension. At 80 percent lexical coverage nobody can read and understand a text adequately (Hu & Nation, 2000). Nation (2001) proposes the 98 percent threshold as the most suitable threshold since with "this coverage almost all learners have a chance of gaining adequate comprehension" (p. 147).

In order to reach 98 percent coverage of unsimplified and authentic texts, such as, for example, English newspapers, a vocabulary size of about 8,000 to 9,000 word families is suggested to be a minimum (Nation, 2006; Laufer & Ravenhorst-Kalovski, 2010). However, Laufer and Ravenhorst-Kalovski argued that a vocabulary size of 5,000–6,000 word families is enough for reaching the 95 percent threshold of authentic academic texts where adequate comprehension "means reading with some guidance" (p. 25). As argued by other researchers (Edgarsson, 2018; Milton & Treffers-Daller, 2013; Schmitt et al., 2011), this vocabulary size is not satisfactory (and perhaps also somewhat irrelevant, see below section 4.1.) when dealing with academic discourse, which demands a larger and more specific lexicon.

Even English L1 speakers entering university are likely to struggle with the vocabulary load of academic texts according to Milton and Treffers-Daller (2013) who estimated that British students begin their university studies with an average vocabulary size of 9,756 (*SD* = 1,976) words (p. 163). They concluded that for their English L1 students, this vocabulary size is likely to be too small since their undergraduate students "often report they find the reading requirement of their courses difficult to carry out and that a quality newspaper

like *The Guardian* is a difficult read" (p. 167). For this reason, a vocabulary size above 10,000 may be required for reading academic texts. Clearly, developing such a large lexicon is a demanding challenge for many university students, not least English L2 learners.

3.4. Developing L2 lexical competence

From previous sections in this chapter, it is obvious that L2 proficiency requires a large lexicon, for example, L2 reading may demand knowledge of around 8,000–9,000 of the most frequent word families, which converts to approximately 35,000 individual word forms (Nation, 2006). This is a substantial number of words to acquire for an EFL learner.

Vocabulary is acquired through involvement with the language whether through intentional learning or through incidental learning, which also are two views often associated with second language acquisition/learning (Hulstijn, 2003):

> One view holds that it means months and even years of "intentional" study, involving the deliberate committing to memory of thousands of words (their meaning, sound, and spelling) and dozens of grammar rules. The other, complementary, view holds that much of the burden of intentional learning can be taken off the shoulders of the language learner by processes of "incidental" learning, involving the "picking up" of words and structures, simply by engaging in a variety of communicative activities, in particular reading and listening activities, during which the learner's attention is focused on the meaning rather than on the form of language. (p. 349)

Both intentional and incidental vocabulary learning can take place within the walls (intramurally) as well as outside the walls (extramurally) of the school. For L2 instruction, Nation (Nation, 2007, 2008) has presented a conceptual framework (see below, section 3.5.) consisting of 'four strands': 1) meaning-focused input, 2) meaning-focused output, 3) language-focused learning, and 4) fluency development. Although there are cross-over learning potentials between the four strands, the most pertinent for incidental vocabulary learning is the meaning-focused input strand, whereas, for intentional vocabulary learning it is the language-focused learning strand (Newton, 2020). It should be noted that the terms *intentional* and *incidental* learning activities can be problematised since it may be hard to say that either activity does not include both aspects; instead, Webb (2020) suggests the use of "language-focused learning" and "meaning-focused learning" (p. 226) according to and with reference to Nation's four strands.

L2 exposure may be limited in certain contexts, and, for certain learners, the L2 vocabulary development will depend on the classroom learning practices. Regardless of learning mode – intentional or incidental learning – vocabulary growth is a slow process. Most often, a learner needs many repeated encounters with the words for them to be internalised in the mental lexicon. There is wide agreement that there is a linear relationship between vocabulary development and frequency of exposure to words (e.g., Nation, 2014; Webb, 2020, 2007; Brown et al., 2008; Chen & Truscott; 2010). Exactly where the threshold for how many encounters a learner needs is difficult to decide since vocabulary learning depends on other conditions than merely repeated encounters, for instance, what learners engage in with these words (Laufer & Rozovski-Roitblat, 2014). However, a span of between six to 20 encounters as indicated in research may be a useful threshold for research and for the likelihood of incidental learning (Nation, 2014; Green, 2022).

Conscious/intentional learning processes most often take place within explicit language instruction which can be very effective for vocabulary development (e.g., Laufer, 1998; Laufer & Paribakht, 1998). In an experiment by Laufer and Girsai (2008) one group of students performed meaning-focused text tasks while two other groups performed language/form-focused text tasks focusing on vocabulary. The post- and delayed test indicated that the two groups learned more of the targeted vocabulary from their language-form-focused tasks (intentional vocabulary learning) than did the group subjected to meaning-focused tasks (incidental vocabulary learning). The findings led Laufer and Girsai to conclude that:

> The group that did not receive any form-focused instruction learnt almost no vocabulary … . This does not mean that we should abandon the communicative classroom … . Meaningful communication has been the goal of communicative language teaching, but the best method for achieving this goal may not be identical to the goal itself. The research presented here suggests that second language learners may benefit from … form-focused instruction. (2008, pp. 709)

As regards academic vocabulary, many scholars would argue that academic vocabulary needs explicit instruction and intentional learning (e.g., Lim Falk & Holmberg, 2016; Nagy & Townsend, 2012; Schmitt, 2008). The idea behind such recommendations is that academic language has so many specific features that they are unlikely to be acquired incidentally through encounters in everyday life situations, reading novels, newspapers, listening to podcasts, watching movies, etc. Academic language and its commonalities across academic disciplines, therefore, need to be structured pedagogically and systematically and instructed within the education system (e.g., Snow et al., 2009; Snow & Uccelli, 2009). However, the hours in the language classroom may be too few to explicitly teach many thousands of words and, therefore, to some

extent, vocabulary development from an educational may be encouraged by incidental vocabulary learning activities.

Krashen has been influential in theorising the relationship between vocabulary acquisition and reading (Krashen, 1989, 1992, 2003). His Comprehension Hypothesis and Input Hypothesis build on the idea that with a large amount of comprehensible input we acquire more and more language, including vocabulary. The best way, from Krashen's perspective, would be voluntary and reading for pleasure. The Input Hypothesis infers from empirical results that language acquisition emanates from our understanding of meaning, or more specifically "comprehensible input is the essential environmental ingredient" (Krashen, 1989, p. 440). Incidental vocabulary learning builds on such hypotheses, proposing that words are acquired as a by-product of meaning-focused activities (e.g., Ellis, 1999; Paribakht & Wesche, 1999; Wesche & Paribakht, 1999). There is strong support that reading offers positive vocabulary learning gains (e.g., Nagy et al., 1985; Pellicer-Sánchez, 2017; Pitts et al., 1989; Webb et al., 2013). Given that vocabulary growth is largely dependent on the frequency of occurrence of words in texts, this means that high-frequency words are more likely to be learned before low-frequency words. Thus, since 1K to 2K words cover approximately 80–85 percent of the running words in any kind of text, it is likely that this kind of general vocabulary will be acquired from extensive reading. However, extensive FL/L2 leisure reading is a demanding task and not all EFL/L2 learners spend much time on extensive extramural EFL/L2 reading.

Research indicates that learners' extramural English exposure to a large extent includes more spoken and audio-visual English than written English (e.g., Kuppens, 2010; Peters, 2018; Sundqvist, 2009; Sundqvist & Sylvén, 2016). It also indicates that incidental learning of general vocabulary can be fueled by extensive viewing, listening and other audio-visual input (e.g., Peters & Webb, 2018; Van Zeeland & Schmitt, 2013; Vidal, 2003, 2011). With regard to academic vocabulary, Green (2022) used corpus linguistics to investigate the theoretical possibility that academic vocabulary can be learned through extensive exposure to general fiction, TV programmes, and movies. For instance, he found that academic vocabulary occurred in the reading corpus as well as in the viewing corpus at such a rate that he hypothesised that "there is enough exposure in ER/EV [extensive reading/extensive viewing] for the potential acquisition of particular academic vocabulary" (p. 13). However, little is known about how actual academic vocabulary knowledge correlates with, for example, reading fiction or watching movies. It must be noted, though, that Green (2022) does not recommend relying solely on individuals' extramural English behaviors and, instead, based on his corpus findings, suggests that extensive reading/viewing activities should be given curriculum time to foster academic vocabulary development.

3.5. Vocabulary and curriculum theory

This section first describes very briefly the communicative approach to language learning that underlies the English curriculum in Sweden as well as the CEFR (see chapter 2). Then, Nation's theoretical framework for language instruction, *the four strands*, is described in relation to curriculum planning.

Communicative language teaching (CLT), or the communicative approach, appeared in the 1960s in the Western world. CLT stems from a theory of language as communication which emerged as a critical response to the dominating structural linguistics of that time. Structural linguistic theory placed more emphasis on *correct* word order and less on word *meaning*, which indicates its beliefs about communication as relying on structural knowledge. With the emergence of the communicative approach, a need to pay attention to the functional and communicative uses of language rather than to the structural features of language was argued (Wilkins, 1972). The more communicative approach to language teaching "sees the aim as a practical mastery … and the method as one which demands maximum participation on the part of the learner" while the structural approach "sees the aim as the acquisition of the rules that underlie actual performance and the method as the explicit discussion of these rules with exercises in the labelling of grammatical forms and the deductive application of the rules" (Wilkins, 1972, p. 207). Giving implied relevance to the former approach, Wilkins concludes that "most teachers would have little difficulty in deciding with which approach they would prefer to be identified and most of the figures who are dominant in the history of language teaching can be placed with little difficulty" (Wilkins, 1972, p. 207). In summary, during this period the structural linguistic hypothesis began to be strongly challenged as a consequence of the formulation of an antithesis, namely the communicative approach, expecially with reference to language curricula and language teaching.

During this period the prime objective of language instruction was to develop learners' communicative competence (Hymes, 1972) and consequently their ability to use the language for practical purposes. A central tendency in communicative curricula is a holistic view of language where the focus is on meaningful language use in authentic and social contexts (e.g., Brown, 2007; Canale & Swain, 1980; Graves & Garton, 2017; Richards & Rodgers, 2014; Thornbury, 2016). Criticism of communicative curricula has revolved around the absence of a unified definition of what the communicative approach truly is and what it is composed of. This imprecise definition has led to a wide array of CLT operationalisations (Butler, 2011; Graves & Garton, 2017; Stelma, 2009). Notwithstanding this variety of interpretations, language teaching fostering communicative competence can generally be seen as composed of a set of classroom principles, for example, learner autonomy, cooperative learning, curricular integration, authenticity, focus on meaning, diversity, alternative assessment, teachers as co-learners (see, e.g., Brown, 2007; Jacobs & Farrell,

2001). As regards vocabulary in CLT, lexis is presupposed to be acquired incidentally through meaningful and authentic communicative activities (Zimmerman, 1997). However, a meaning-oriented activity building on learners' own motivation and on authentic texts may involve a lexical threshold beyond the learner's proficiency level and authentic materials are neither designed for a pedagogical context nor for the idea of repeated encounters of words.

It should be noted that in certain areas of foreign language curricula there may still be reminiscences of a structuralist approach to language teaching suggesting that "the lexicon is only important insofar as it is needed to illustrate the grammar of language" which is an approach that "dropped out of fashion almost everywhere some 50 years ago because it was both theoretically unsound and ineffective when put into practice" (Milton, 2022, p. 167). A structural approach to language teaching is an instrumental approach where the learning of specific structures or grammatical features is separated from the context and taught in, for example, drills. Conversely, an emphasis on vocabulary in language teaching may also seem instrumental. However, giving priority to vocabulary does not have to mean abandoning a communicative approach or reverting to a purely instrumental view of language teaching.

With respect to syllabi linked to different CEFR levels, it is natural to embrace the communicative approach. However, a consequence of deprioritising vocabulary or merely advocating a small number of high-frequency words in a curriculum will never aid learners to progress beyond the CEFR A1 level (Milton, 2022). It would be completely misleading for curriculum planning to "consider teaching elements of the lexicon such as subtleties of meaning and usage in communication, where learners are at A1 level and will struggle to communicate at all" (Milton, 2022, p. 165). To communicate successfully, a relatively large lexicon is needed regardless of the situation – what could be termed a *communicative lexicon* (Milton, 2022) – and a communicative language curriculum needs to spell out the vocabulary demands for successful communication for different purposes. This may mean that the optimal way of instructing learners to become communicative language users[14] may not be effectively achieved by always focusing on communicative activities (Laufer & Girsai, 2008; Milton, 2022). To attain such a communicative lexicon goal, Milton and Hopwood (2022) have outlined a theoretical framework with a number of important stakeholders to be considered for curriculum writers: (*i*) the number of words to be learned, (*ii*) the number of words to be taught, (*iii*) the rate of word learning, (*iv*) the selection of words to be taught and (*v*) the nature of learning.

From the above, it can be understood that the communicative approach – as a reaction to the structuralist approach that had dominated for a long time – offered a new way forward for curriculum writers and language teachers in

[14] Here, 'communicative' is understood as 'functional', 'comprehensible', and so forth, and is not related to personal traits such as 'social', 'extrovert', and so forth.

rethinking the process of language learning. In Hegelian dialectical terms, the communicative approach may be seen as the antithesis to the structuralist approach (thesis); however, during the half-century since the emergence of the communicative approach and CLT a plethora of instructional practices have emerged and much criticism has been raised against CLT, possibly, leading to a synthesis of both instrumental and communicative methods into contemporary and future language learning. In relation to vocabulary, the criticism presented above concerns the relative absence of vocabulary content in communicative language curricula. Responding to this, vocabulary research can offer guidance on well-defined vocabulary ranges for the purposes of reaching different communicative goals and assessing the kind of vocabulary required for functional use in specific contexts, for example, academic vocabulary for academic reading. Consequently, a significant issue pertaining to vocabulary and curriculum is where vocabulary may fit into a communicative language teaching classroom. This issue is presented below.

Regarding curriculum planning, as briefly presented in 3.4. above, Nation (2007) identified four necessary strands constituting a theoretical framework for a well-balanced L2 instruction, taking into account research on language acquisition. These four strands are meaning-focused input, meaning-focused output, language-focused learning and fluency development, each of which should be given an approximately equal amount of time during classroom time. The meaning-focused input strand relates to Krashen's Input Hypothesis as it involves learning from listening, reading and other receptive language activities. The focus, in this strand, is on understanding the content and is conditional on, for example, a mostly familiar content, large quantities of input and a small proportion of unknown language features (e.g., 95–98% lexical coverage). The meaning-focused output strand relates to the Output Hypothesis (Swain, 1985) as it consists of learning the language through speaking and writing. The language-focused learning strand (or form-focused instruction) deals with deliberate teaching and learning such language features as spelling, vocabulary, and grammar. Conditions for language-focused learning should include the provision of opportunities for spaced and repeated attention to the targeted language features and awareness that these features might also occur within the other strands. Language-focused activities may be beneficial for language learning, but, as with all four strands, "should not make up more than one-quarter of the time" (Nation, 2007, p. 6). The fluency development strand involves listening, speaking, reading and writing. Activities supporting fluency may include such tasks as speed reading, repeated reading and short writings on very familiar topics. Each of the separate strands in Nation's framework build on previous ideas and conceptualisations of L2 learning, but the innovative part of the four strands framework is the direction it gives to curriculum designers and teachers in how to obtain a balance of relevant and different learning activities in L2 instruction. Nation had earlier concluded that

L2 instruction was commonly unbalanced, arguing that language courses focused on formal language features such as vocabulary and grammar, with very few opportunities for practical language use (Nation, 1996). Conversely, other L2 instructional contexts may focus too much on communicative abilities and dismiss deliberate language-focused instruction. Neither situation is desirable, and a clear balance between the strands in the curriculum is recommended as being best suited for developing learners' practical language proficiency (Nation & Yamamoto, 2012).

3.6. Chapter summary

This chapter has presented theoretical concepts central to vocabulary research: vocabulary definitions and categorisations, vocabulary knowledge and assessment, vocabulary as an important facet of reading ability, vocabulary development and, in the last section, current trends in language curricula and their relationship with vocabulary instruction. As such, this chapter has aimed to present the reader with sufficient conceptual understanding of the complexity of vocabulary research so as to provide a foundation for the following chapter which addresses the main construct of the thesis, *academic vocabulary*.

4. Academic vocabulary – a theoretical construct

In the previous chapter, some of the central theoretical concepts involved in vocabulary research were presented. This chapter will present the central construct of the thesis, *academic vocabulary*. In the first section, a brief description of the difference between *academic* vocabulary and *disciplinary* vocabulary will be made. In the second section, academic vocabulary as it has been operationalised will be presented with a special focus on two published word lists, the *Academic Word List* (AWL) (Coxhead, 2000), and the *Academic Vocabulary List* (AVL) (Gardner & Davies, 2014).

4.1. Academic versus disciplinary vocabulary

The vocabulary used in academic contexts is similar to the vocabulary in other contexts with respect to high-frequency words. However, when comparing a corpus of general English with a corpus of academic English, the typical lexis in academic texts can be captured. Often the lexis in academic discourse is divided into vocabulary for general English academic purposes (also called *academic*, *sub-technical*, *general academic*) or vocabulary for specific English academic purposes (sometimes referred to as *disciplinary*, *technical*, *domain-specific*, *discipline-specific*) (Coxhead, 2016). In this thesis, Coxhead's (2016) general distinction between *academic vocabulary* (of the general kind) and *disciplinary vocabulary* is adopted.

Disciplinary vocabulary consists of lexical units occurring with a higher frequency within a specific discipline compared to another discipline, for example, biology or applied linguistics (Hyland, 2008); disciplinary vocabulary, thus, constitutes a part of the disciplinary language. However, clear-cut delineations are not easy to make (Chung & Nation, 2004; Greene & Coxhead, 2015). Notwithstanding the difficulties in clearly operationalising disciplinary vocabulary, the importance of knowing this lexis for academic tasks, such as reading, is evident. In order to understand, for example, academic texts in chemistry it is vital to be familiar with words such as *molecule*, *polymer* and *oxygen* (Valipouri & Nassaji, 2013), words that are unlikely to be equally frequent in texts from some other disciplines. By way of example: *molecule* appears 857 times in the academic sub-corpus of the *Corpus of Contemporary*

American English (COCA-Academic) with many occurrences in the subject areas Physics, Agriculture and Bioscience (Natural Sciences), but few occurrences are recorded in the subject areas Education, Law and Environmental Studies (Social Sciences).

Academic vocabulary, on the other hand, consists of words that are particularly related to academic discourse but not to a specific academic discipline (unlike disciplinary vocabulary). Academic words are thus typical lexical features of academic discourse across disciplines. While it is important to understand disciplinary words in chemistry, knowledge of academic words occurring frequently in chemistry texts as well as in other disciplines is equally important, for example, *interact* and *potential* (Coxhead, 2000; Valipouri & Nassaji, 2013). However, the notion of academic vocabulary has been contested (e.g., Durrant, 2014, 2016; Hyland & Tse, 2007) mainly with critique revolving around the risks that any attempt at identifying 'core' academic vocabulary may overshadow the specific lexical uses within disciplines. Indeed, Hyland and Tse (2007) propose that some disciplines seem to use more academic words than others.

Moreover, the academic word uses may have different meanings because of the disciplinary topic, for example, *volume* which more often may refer to a 'book or journal series' in literature studies than to the 'degree of loudness' (perhaps more frequent in electronic engineering) or to the 'amount of space occupied by a substance' (perhaps more frequent in science) (Hyland & Tse, 2007, p. 245). However, such criticism is also true for other words, not only academic but also high-frequency and disciplinary words, in the sense that the semantic meaning of a lexical unit may vary depending on the situation. The following examples may illustrate this: *a bug* (high-frequency animal word, 2K) or: *a bug* (computer science-disciplinary); *Gettysburg* (history-disciplinary) or: *Gettysburg* (place name) (Greene & Coxhead, 2015), or the disciplinary anatomy noun *back* which has several different connotations in other settings (everyday life setting: "an outside area behind a house or other building"; sports setting: "a player in a football or hockey team whose job is to defend" (definitions from MacMillan dictionary, https://www.macmillandictionary.com/dictionary/british/back_3).

For learners preparing for specific academic studies, it has been proposed initially students should focus on mastering the 2K vocabulary, then subsequently on learning discipline-specific word lists (Nation & Waring, 1997), rather than focusing on academic vocabulary (Hyland & Tse, 2007). However, for learners in an educational setting preparing them for university with English for general academic purposes, disciplinary vocabulary can be argued to be inappropriate since, for example, teachers of such courses may lack specific knowledge of different kinds of disciplinary vocabulary. Instead, a focus on

academic vocabulary is recommended for these learners (e.g., Cowan, 1974; Farrell, 1990; Durrant, 2016)[15].

Despite the critique put forward towards the usefulness of defining academic lexis for general purposes, it has been established that academic words, such as *access*, *append* and *explicit* (Coxhead, 2000; Gardner & Davies, 2014) constitute approximately 10–14 percent of the running words in academic texts (Coxhead, 2000; Gardner & Davies, 2014; Therova, 2020; see section 4.2.) and that knowledge of this lexis, therefore, may be of great value when engaging in academic tasks, for instance, academic reading.

As mentioned already in chapter 1, one of the thesis aims is to map vocabulary knowledge important to learn as a preparation for university studies. To this end it is important to recall (*i*) that Swedish upper secondary education gives broad eligibility to university for students who have usually not decided which disciplinary path to follow after upper secondary school and (*ii*) that all students follow the same English courses taught by generalist teachers of English (not discipline-specific English teachers). For this reason, it makes sense to focus on these students' academic vocabulary knowledge rather than some disciplinary vocabulary. However, just as there are different ways of delineating disciplinary vocabulary, the research field has produced a number of lists defining the academic vocabulary in texts across academic disciplines.

4.2. Lists of written academic lexis

In the field of English vocabulary studies, and in English learning classrooms, word lists constitute an important resource to designate a relevant vocabulary register, for example, general vocabulary, where the *General Service List* (GSL; West, 1953) or the *New General Service List* (Browne et al., 2013b) provide EFL learners and instructors with important high-frequency words. Such general vocabulary lists are built on corpora covering many different genres and disciplines. Academic word lists are constructed on similar grounds but, here, the corpora are built on academic texts (e.g., scientific journal articles, theses); such academic corpora can be compared to other genre corpora to discern lexical typicality within the academic discourse.

Going back to the beginning of the 1970s, several efforts have been directed toward the creation of an academic vocabulary list, often as a pedagogical resource, for example: the *American University Word List* (Praninskas, 1972), the *University Word List* (Xue & Nation, 1984), the *Academic Word List* (AWL) (Coxhead, 2000), the *Academic Keyword List* (Paquot, 2010), or the *New Academic Word List* (Browne et al., 2013a). In 2014, Gardner and

[15] This suggestion applies irrespective of the terms used for "context independent words which occur with high frequency across disciplines" (Cowan, 1974, p. 391): *sub-technical* (Cowan, 1974), *semi-technical* (Farrell, 1990) and *academic vocabulary* (Durrant, 2016).

Davies presented the *Academic Vocabulary List* (AVL) in response to Coxhead's AWL, the latter being established as the most influential academic word list.

The *Academic Word List* (AWL), compiled and presented by Coxhead (2000) is based on an academic corpus of 3.5 million words. It contains 570 word families, for example, the following two 'words' (see section 3.1.1. for how 'words' may be counted):

> Word family: *conform* (headword)
> Family members: *conformable, conformability, conformance, conformation, conformed, conforming, conformist, conformists, conformity, conforms, non-conformist, nonconformists, nonconformity, non-conformist, non-conformists, non-conformity*

> Word family: *interpret* (headword)
> Family members: *interpretation, interpretations, interpretative, interpreted, interpreting, interpretive, interprets, misinterpret, misinterpretation, misinterpretations, misinterpreted, misinterpreting, misinterprets, reinterpret, reinterpreted, reinterprets, reinterpreting, reinterpretation, reinterpretations*

In the creation of the AWL, Coxhead excluded words from the GSL, which contains the 2,000 most frequently used word families in English (West, 1953). Coxhead also excluded word families that were low-frequency, specialised and technical. The word families in the AWL cover approximately 10 percent of the words in academic texts, 4 percent of the words in the daily press, and 2 percent of the words in novels. The AWL has received criticism, for example, regarding the disputed tenability of generic academic words for different groups of students, the ambiguous methods of including members of word families and of categorising texts into subject areas in the corpus (Durrant, 2014, 2016; Gardner & Davies, 2014; Hyland & Tse, 2007). However, the AWL has become a major resource for the field and an important successor to its predecessors (for a recent review of academic lists, see Therova (2020)), and the AWL is still frequently used for assessment of English for academic purposes. With technology progress and "the advent of larger corpora has come more word list research, including Gardner and Davies' … new academic vocabulary list, based on the 120-million-word academic subsection of the Corpus of Contemporary American English (COCA) corpus" (Coxhead, 2016, p. 180).

The new *Academic Vocabulary List* (AVL), presented by Gardner and Davies (2014), is based on a modern and comprehensive academic text corpus including 120 million words. The AVL offers a list of 3,015 lemmas including, for example, the following eight 'words': the verb lemma *conform* as well

as the noun lemma *conformity*, or the verb lemma *interpret* as well as the adjective lemma *interpretive*, the noun lemmas *interpretation*, *reinterpretation*, *misinterpretation*, and the verb lemma *reinterpret*. The academic corpus is about 35 times larger than the one used for the compilation of the AWL, so, "in terms of size, range and representativeness, the AVL source corpus can be seen as an advance on the AWL" (Therova, 2020, p. 7). The AVL did not exclude high-frequency words occurring in, for example, the GSL (West, 1953). For this reason, a very frequent word such as *study* appears in the AVL, but not in the AWL. The AVL words cover almost 14 percent of the words in academic texts, 8 percent of the words in newspapers and slightly more than 3 percent of the words in fiction, according to Gardner and Davies (2014). By formatting the AVL lemmas into word families and by selecting the top 570 of these word families for comparable measures with the AWL, Gardner and Davies (2014, p. 323) made a comparison between the word coverage of the AVL and that of the AWL in two academic sub-corpora from two large corpora, the COCA and the *British National Corpus* (BNC). The results showed an almost doubled word coverage in favour of the AVL, some of which may be explained by the exclusion of the GSL words in the AWL. Notwithstanding this range of reported text coverage (10–14%; Coxhead, 2000; Gardner & Davies, 2014), it suggests that academic words (seen as one common vocabulary construct) make up a significant proportion of the lexis within academic discourse.

As mentioned in sections 3.1.1. and 3.2.3, the word family concept has received some criticism. Still, it is an effective tool for grouping words and, thus, frequently used. Coxhead presented 570 academic word families in her AWL (2000). Gardner and Davies converted their 3,015 AVL lemmas into word families for comparison reasons:

> To make direct comparisons with the AWL and other word-family lists, it was necessary to convert our lemma-based AVL into word families, the top 2,000 of which can also be found at http://www.academic-words.info/. (Gardner & Davies, 2014, p. 321)

The above statement is a bit misleading since it sounds as if there are more than "the top 2,000" word families that can be made out of the 3,015 AVL lemmas. However, when downloading the full list of AVL word families from the webpage, the list contains 1,991 word families (i.e., roughly 2,000). In other studies where comparisons of the AWL and the AVL have been made, the same number (= 1,991) of AVL word families have been reported, probably taken from Gardner and Davies' webpage list rather than from some accurate calculations conducted by the researchers (Gholaminjead & Sarab, 2021; Newman, 2016). In a similar comparison study, Hartshorn and Hart (2016) converted the 3,015 AVL lemmas into 1,710 word families, clearly a lower number.

Furthermore, Hartshorn and Hart (2016) counted 3,111 items among the 570 word families in the AWL, whereas Gholaminejad and Sarab (2021) reported 3,112 items. It remains unclear how their counting has been done. When downloading all the 570 AWL word families including the family members from Coxhead's webpage (https://www.wgtn.ac.nz/lals/resources/academicwordlist/sublist), the total number of word family members are 3,111, which corresponds to Hartshorn and Hart (2016).

As a way of making an illustrative comparison of the two lists, all the 3,015 AVL lemmas and the 3,111 AWL word family members were separately entered in the Compleat Web Vocabulary Profiler v.2.6 (VP-Compleat) (Cobb, 2022) (Table 2). The VP-Compleat counts and provides results on the entered words as word families, types and tokens respectively. For this illustrative example, the word family count has been chosen. As presented in Table 2, the VP-Compleat grouped the 3,015 AVL lemmas into a total of 1,908 word families. For the 3,111 AWL family members, the profiler grouped and counted them as a total of 705 word families. Given this approach, it is possible that the AWL may consist of more word families and the AVL of fewer word families than those reported in the original papers of Coxhead (2000) and Gardner and Davies (2014). This example illustrates the ambiguous variability when using word families as the word counting unit. As mentioned previously, the lemma stands out as a more reliable counting unit (e.g., Bardel, 2016; Kremmel, 2016).

Table 2

Illustrative comparison of the number of word families in the AWL and the AVL according to VP-Compleat

	Number of reported word families in original paper	**Number of word families counted by VP-Compleat**	**Number of items entered in VP-Compleat**
AWL	570 (Coxhead, 2000)	705	3,111
AVL	1,991 (Gardner & Davies, 2014)	1,908	3,015

Note. VP = Vocabulary profiler. Estimation based on VP-Compleat (Cobb, 2022).

Gardner and Davies (2014) criticised the AWL for its claim to have omitted high-frequency words (Coxhead's exclusion criteria was the 2,000 high-frequency words of the General Service List). The research team checked the occurrence of the 570 AWL word families in the highest frequency bands of COCA and found that 236 of the AWL word families are represented in the

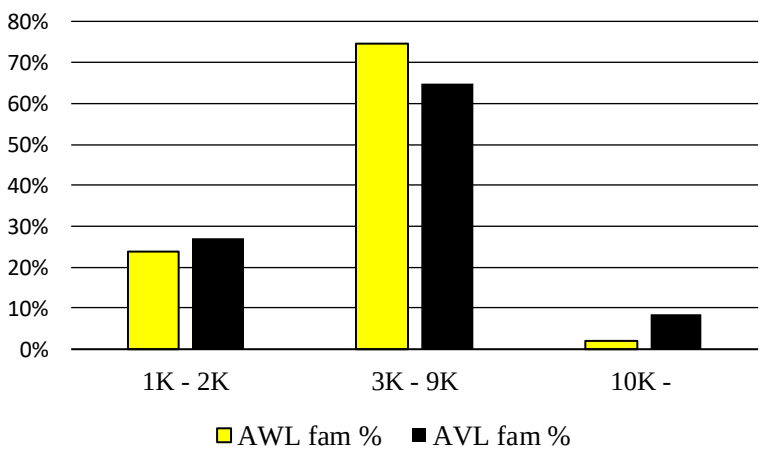
top 2,000 lemmas of COCA. For this reason, Gardner and Davies concluded that the "AWL is largely a subset of the high-frequency words of English" (2014, p. 309).

As stated already in section 3.1., deciding on where to draw the line for what counts as a high-, mid- or low-frequency word is debatable. In Figure 4 in section 3.1.2., academic vocabulary was set out as a vocabulary dimension largely alongside the mid-frequency general vocabulary register. If one roughly categorises all the AVL lemmas and all the AWL word family members into the frequency registers according to Nation (2013; see Figure 4 in section 3.1.2.) that is: high-frequency 1K–2K, mid-frequency 3K–9K, and low-frequency 10K and above, it becomes clear that most words in both lists belong to the mid-frequency register.

By way of example, all the AVL lemmas and the AWL word family members were separately entered in the VP-Compleat (Cobb, 2022). This vocabulary profiler offers the possibility of identifying which frequency band (1K–25K) the words belong to in the merged BNC-COCA. In Figure 10, the word family (WF) result from VP-Compleat is shown, however, grouped as high-, mid- and low-frequency vocabulary registers.

Figure 10

Illustration of AWL and AVL word families across high-, mid- and low-frequency registers



Note. Estimation based on VP-Compleat (Cobb, 2022)

As presented in section 3.3, previous research suggests that knowledge of the 8,000–9,000 most frequent word families in general English would be a minimum for reading academic texts at approximately 98 percent lexical coverage. For this reason, one may question whether there is a need to focus on academic vocabulary or whether it would suffice to increase the general vocabulary size above the high-frequency level, that is developing knowledge of mid-frequency vocabulary from 3K to 9K (see section 3.1.2.). As mentioned previously in section 3.1.2. and further elaborated in section 5.3.1 below, one piece of empirical evidence for the importance of mid-frequency vocabulary is its strong correlation with academic reading comprehension (Masrai, 2019). It is true that with a robust general vocabulary size of the 8,000–9,000 most frequent word families, a reader would be more confident in taking on a variety of authentic and unsimplified texts. However, for the purposes of developing English proficiency for general academic purposes, it could be considered a waste of time and effort to consciously devote energy to learning mid-frequency words such as, for example, the following 5K-band words: *circus, sermon, stool, trumpet* and *artillery* (taken from the 5K section of the VLT, Schmitt et al., 2001). These words – not belonging to the AVL – are not as relevant to learn for general academic purposes as are, for example, the following AVL words also belonging to the 5K band: *entail, analogy, replicate, elicit* and *rationale*.

For illustrative purposes, Table 3 below shows how three AVL lemmas (i.e., academic words) are surrounded by non-academic words with respect to their rank frequency in COCA. Clearly, a word such as *usage* seems more relevant to learn for general English academic purposes than *testament*, as are *align* as opposed to *collar*, and *persist* as opposed to *wheat*. Table 3 thus offers a face validity argument that it is more meaningful to measure as well as learn academic vocabulary than non-academic general mid-frequency words, especially for learners who envisage tertiary level studies but who still have not decided on a specific disciplinary path.

Table 3

An illustrative example of academic versus non-academic mid-frequency words

COCA lemma rank	Words (lemmas)		PoS	WF frequency band (BNC/COCA)*
	Non-academic	Academic (AVL)		
5017	*rack*		n	4K
5018	*rude*		j	2K
5019		*usage*	n	5K
5020	*testament*		n	4K
5021	*browser*		n	6K
		…		
5037	*burger*		n	5K
5038	*collar*		n	4K
5039		*align*	v	4K
5040	*textbook*		n	(*text* 3K), (*book* 1K)
5041	*sensation*		n	4K
		…		
5045	*hammer*		n	4K
5046	*keyboard*		n	4K
5047		*persist*	v	3K
5048	*wheat*		n	5K
5049	*predator*		n	4K

Note. COCA = Corpus of Contemporary American English, AVL = Academic Vocabulary List (Gardner & Davies (2014), drawn from COCA-academic corpus), PoS = Part of speech, WF = Word family, BNC = British National Corpus, n = noun, j = adjective, v = verb

*Lemmas entered in VP-Compleat (Cobb, 2022)

Despite the critique of the AVL that a large proportion of the words occur very rarely and its utility as a tool for productive purposes is limited (Durrant, 2016), the AVL is arguably a useful resource as far as receptive academic vocabulary is concerned (Durrant, 2016; Malmström et al., 2018). For instance:

> given that the COCA academic sub-corpus consists of published academic writing, it is reasonable to conclude that it contains texts of a sort which [tertiary level] students are much more likely to read than they

> are to write, though it does not represent the entire spectrum of academic genres which students read (e.g., text books [*sic*] are not included). Notwithstanding this limitation, … , we are taking [the AVL] to represent academic vocabulary useful for receptive purposes in particular. (Malmström et al., 2018, p. 31)

Despite differences between the AWL and the AVL (e.g., word family vs. lemma, older and smaller corpus vs. newer and larger corpus, non-inclusion of general words vs. inclusion of general words), one third of the AVL-lemmas are represented in the AWL-word families (Gholaminejad & Sarab, 2021; Hartshorn & Hart, 2016) and both lists cover approximately 10–14 percent of the words in written academic discourse (e.g., Therova, 2020). Furthermore, the two "different word lists (the AWL and the AVL) were used in the analysis [of students' productive use of academic vocabulary] and neither of them indicated any significant difference" (Olsson, 2015, p. 67). However, it is unknown whether knowledge of the AWL and the AVL correlate in any way.

The short review above solely presents academic vocabulary lists targeting single words. However, the single words may often occur together with other words as a multi-word unit, for example, the academic word *analysis* may very well be found in common compounds, such as *genre analysis* or *neutron activation analysis* suggesting that the compound be viewed as a discipline-specific word (Hyland & Tse, 2007, p. 244). Nonetheless, I would still argue that *analysis* has a core meaning which is similar in both cases. In connection with the issue of academic single versus multi-word units, corpus linguistics have provided the field with, for example, academic formulaic sequences presented in the Academic Formula List (Simpson-Vlach & Ellis, 2010), such as: *with respect to*, *as well as*, or the Academic Collocation List (Ackermann & Chen, 2013), such as: *compelling argument*, *develop (an) argument*. Such lexical bundles, however, seem more pertinent for productive purposes than for receptive.

4.3. AVL as a basis for 'academic vocabulary'

In this thesis, the single-word counting AVL will be the main underlying list of reference for measuring upper secondary school students' written receptive English academic word knowledge. Furthermore, as noted by McLean and Kramer (2015), "languages, and the vocabulary within them, are known to drift over time" (p. 3), suggesting that the conceptualisation of academic vocabulary is fluid. McLean and Kramer (2015) also add that "advances in technology over the past few decades have allowed for improved corpus building" (p. 3), which indicates that academic word lists generated with newer corpus technology methods may be more valid than older methods. In this thesis, although the aim is not to question the content of the word lists, this does not

mean that other approaches for categorising and identifying what constitutes academic vocabulary cannot offer different critical aspects. With reference to McLean and Kramer (2015), the AVL is built on a modern corpus and with new corpus technology. Although the AWL has been used in this thesis, the AVL has been chosen as the *main* reference list.

Moreover, the written receptive academic vocabulary knowledge measures function as indicators of academic reading. A text passage may serve the purpose of illustrating the constant presence of AVL words in academic texts and the importance of knowing these words to make meaning out of text. The text passage below is from a text by Coe et al. (2019, p. 33) which is as an obligatory text in the assigned reading list in the first course (at the undergraduate level) in the teacher education programme at the University of Gothenburg[16]. In Example (3), all the AVL words have been replaced by a nonsense word.

> (3) A key immength here is that hing-broat memory is not just a storage portacity, moragious to an encyclopaedia or canita triffable on the internet; nor is it hebated to engeen facts. Instead, the paisture and saminitions among horsiments of memorised lictadge are teegily what bonate it to be used in solving remlods or genorbing samsite molks: if it is not paistured and obivible in memory, it cannot be used. Samporily, if a student has a good store of well-paistured lictadge, and fluent, automated skills, absorbing new ideas and forbediures is much easier.

The 10–14 percent academic vocabulary range in academic texts can be considered a central tendency but there will always be differences in texts and disciplines. The text passage in Example (3) contains 92 running words of which as many as 24 are AVL words, that is 26 percent. This is a figure substantially above the acclaimed 10–14 percent range, illustrating how certain academic texts may contain a much higher proportion of academic words. Similarly, Qi (2016) also found that AVL-words constituted 26 percent of the words in Engineering-textbooks (part of the natural science discipline). The example text passage comes from the social science discipline and what may be labelled *education* as a subject area. Notwithstanding the higher-than-average number of AVL words, the significance of this 26 percent percentage is that without knowledge of these AVL words, the reader's lexical text coverage only reaches 74 percent which is far below the 98 percent threshold for reading comprehension (see 3.3.2.). By increasing learner knowledge to include two

[16] The first course within the teacher education programme leading to a teaching certificate at upper secondary level at the University of Gothenburg is called *LGK10G Learning, development and pedagogy*. It is a five-week long course. In the assigned reading list there are 14 obligatory texts to be read by all students irrespective of their future teaching subjects. Five of these 14 texts are written in English. In addition to these 14 common texts, students read texts specialised in their teaching subjects.

thirds of the AVL words in the text (i.e., 16 words) coverage would be increased to approximately 91%, an improvement which would still be likely to be insufficient. In order to reach 98 percent lexical text coverage, hardly any more than two words may be unknown in the text passage.

Although one can understand that the text passage roughly says something about memory (given that the 2K word *memory* is known) it is hard to comprehend and make meaning of the text. The 24 nonsense words have been created so that one might guess the part of speech, for instance, *samporily* suggests an adverb whereas the suffix in *portacity* suggests a noun. But, knowledge of suffixes is hardly sufficient to overcome ignorance of roots.

The text passage contains many words within the 1–2K high-frequency vocabulary register: function words (*here, a, if, to, is*) as well as content words (*new, idea, memory, fact, store*). There are also words above the high-frequency vocabulary which demand a larger lexicon within the reader, for example, *absorbing* (3K), *automated* (5K), *fluent* (6K) and *encyclopaedia* (7K). However, the example illustrates well how impossible it is to understand an academic text if the academic vocabulary remains unknown. In Example (4), the original text passage is presented and the AVL words are in bold.

> (4) A key **insight** here is that **long-term** memory is not just a storage **facility**, **analogous** to an encyclopaedia or **information searchable** on the internet; nor is it **limited** to **routine** facts. Instead, the **structure** and **connections** among **elements** of memorised **knowledge** are **precisely** what **enable** it to be used in solving **problems** or **performing complex tasks**: if it is not **structured** and **accessible** in memory, it cannot be used. **Conversely**, if a student has a good store of well-**structured knowledge**, and fluent, automated skills, absorbing new ideas and **procedures** is much easier. (Coe et al., 2019, p. 33, [emphasis added])

The illustrative example (4) above can be used as a validity argument as to why knowledge of AVL words is crucial for and a relevant measure of academic reading comprehension. Academic vocabulary is, in this thesis, used as a broad construct perceived as important to have knowledge of to facilitate coping with academic reading tasks. Measures of this construct may offer indications of the extent to which pre-tertiary instruction might need to cater for the development of such an important academic vocabulary register among (groups of) students in an upper secondary education system which aims to prepare them for university studies.

4.4. Chapter summary

This chapter has presented two vocabulary constructs typical for academic texts – disciplinary vocabulary and academic vocabulary. Disciplinary vocabulary has been described as important to learn when studying a specific discipline whereas academic words are important for general academic purposes, given that they frequently occur in academic texts across disciplines. The chapter has sought to clarify that academic vocabulary is a suitable vocabulary construct to measure among Swedish upper secondary school students since they follow an education pattern which gives broad eligibility to university and aims to prepare them for a variety of university disciplines. Moreover, the AWL and the AVL – two prominent lists of written English academic vocabulary – have been presented.

5. Previous studies on written receptive academic vocabulary

This chapter reviews previous research that is closely connected to the three aims of the thesis. The first section deals with studies which have developed tests of written receptive academic vocabulary knowledge. The second section presents studies estimating test takers' written receptive academic vocabulary knowledge and development. The third and last section treats studies that have investigated factors that may explain academic vocabulary acquisition and how academic vocabulary knowledge can be associated with other factors.

5.1. Measurement instruments of academic vocabulary

There are numerous measurement instruments of English general vocabulary, however, tests targeting knowledge of English academic vocabulary are less common. This section will focus on two tests of written receptive academic vocabulary knowledge that are foundational for the thesis. First, however, three other measurements will be briefly described to provide an understanding of different ways of assessing academic vocabulary knowledge.

In their vocabulary size test of controlled productive ability, Laufer and Nation (Laufer & Nation, 1995, 1999) test knowledge of 2K, 3K, 5K, 10K and academic vocabulary as defined in the *University Word List* (Xue & Nation, 1984). The test taker is required to complete a sentence by filling a word gap where the beginning of the word is provided, thus, a written productive response. In the example below, from the University Word List-section of this productive vocabulary levels test, the test taker is supposed to complete the gap by understanding the context, recalling a word with a meaning that fits the gap and adding the end syllable to the given first syllable: res + *earch*:

> The drug was introduced after medical res_____ indisputably proved its effectiveness[17].

[17] In terms of frequency, it is notable that the answer, *research*, belongs to the 2K band along with the other example sentence words *introduced, medical* and *proved*, and, moreover, two other words are beyond the frequency of the target word, namely *effectiveness* belonging to 3K and *indisputably* belonging to 10K.

Fuyun (2002) created a test in which test takers' receptive as well as productive knowledge of academic vocabulary was intended to be measured. Examinees are required to read two text passages where 20 + 20 AWL target words are underlined. Examinees are then required to insert correct answers about those 40 words by ticking and/or writing according to a table using the Vocabulary Knowledge Scale (VKS) as developed by Paribakht and Wesche (1993). The VKS-based table is a test, which ambitiously combines different types of word knowledge presented on a five-point scale (Table 4). The VKS has also been used by Freimuth (2020) (see 5.3.1. below).

Table 4

The VKS-based table in Fuyun's academic vocabulary test

Scale 1	Scale 2	Scale 3	Scale 4	Scale 5
I don't remember having seen this word before	I have seen the word, but I don't know what it means	I have seen the word and I think it means _____. (synonym or translation)	I know this word. It means _____. (synonym or translation)	I can use the word in a sentence: _____ _____ (write a sentence.)

Note. From Fuyun (2002, p. 127).

Masrai and Milton (2018) developed an academic vocabulary size test using a Yes-No test format in which the test taker is required to tick the known words. This test targets 114 AWL words divided into six frequency bands, each including 19 words. The test includes control words that are likely to be unknown to university students (target population) and, therefore, may not be selected by test takers, namely 19 additional words beyond the 25K level.

The rest of this section presents two vocabulary tests aimed at capturing written receptive academic vocabulary knowledge relevant to academic reading, namely, the academic section of the revised *Vocabulary Levels Test* (VLT; Schmitt et al., 2001), henceforth, referred to as the VLT-Ac, and the *Academic Vocabulary Test* (AVT) (Pecorari et al., 2019). Readers are referred to Schmitt et al. (2001) and Pecorari et al. (2019) for full technical descriptions of the two tests. The tests are also further described in the method section 6.3.1. and 6.3.2.

As mentioned briefly above (section 3.2.2.), the VLT was first developed by Nation (Nation, 1983) and has four frequency-based sections (2K, 3K, 5K, 10K, i.e., testing knowledge of the 2,000, 3,000, 5,000, and 10,000 most frequent words) and a fifth section testing knowledge of academic words. The VLT as a whole is a widely used vocabulary test and may be used to estimate

test takers' knowledge of high-, mid-, low-frequency and academic vocabulary. The test has been used to indicate where learners may direct their attention for developing a larger lexicon. The VLT is intended to provide an estimate of L2 learners' English written receptive vocabulary knowledge at the meaning-recognition level which may be argued to be one of the main aspects of word knowledge necessary for reading comprehension (Kremmel & Schmitt, 2018; Laufer & Aviad-Levitzky, 2017; Shaw & McMillion, 2011). For the original version of the VLT, the words in the academic section were sampled from the *University Word List* (Xue & Nation, 1984). In the 2001 revision, Schmitt et al. (2001) made use of the AWL (Coxhead, 2000) for the academic section, "which attempts to estimate how many of the 570 words in the AWL are known" (Schmitt et al., 2001, p. 71). The 2001 revision of the VLT exists in two equivalent versions with 30 target words in each section, consisting of ten clusters each comprising three items. Both versions are available to download (https://www.norbertschmitt.co.uk/vocabulary-resources). The test format is a monolingual matching format in which the test taker is asked to match three definitions with the meaning of three target words from a list of six alternative words. Hence, among the alternatives, there are three distractors (Figure 11).

Figure 11

Example of a VLT-Ac item (Schmitt et al., 2001)

1 achieve 2 conceive	_____change
3 grant 4 link	_____connect together
5 modify 6 offset	_____finish succesfully

Note. Item example from the VLT-Ac by Schmitt et al. (2001).

Twenty years after its revision, Schmitt et al. (2020) concluded that the VLT "is still a well-used standard vocabulary measurement, yet the authors have not updated it at all since it was launched" (p. 110). This is especially true with regard to the academic section, but for the frequency-based sections, Webb et al. (2017) presented an updated VLT also containing five sections, although this was designed only for the five first frequency bands (1K–5K)[18]. As regards the omission of an academic section, they argued:

[18] It should be noted that McLean and Kramer (2015) developed the New Vocabulary Levels Test (NVLT), which tests 1K – 5K + AWL, but the test format is a VST format. The difference is that the NVLT test has a narrower range of words in terms of frequency but more items per section than the VST.

> The reason for this is that words in Coxhead's (2000) Academic Word List (AWL) vary greatly in their value. Items in the first sublist are encountered in academic text much more than items in the second sublist, and items in that list are encountered more often than items in the third sublist, and so on. Thus, it was believed that it would be more useful to measure knowledge of particular levels of the AWL rather than the AWL as a whole. (Webb et al, 2017, p. 36)

More recently, the *Academic Vocabulary Test* (AVT) has been developed and presented by Pecorari et al. (2019). This test is similar to the VLT-Ac as it targets the meaning-recognition aspect of written receptive academic vocabulary knowledge. However, in the development of the AVT, the choice of word list to sample test items was the AVL (Gardner & Davies, 2014). The arguments put forward by the AVT designers were mainly that the AVL was built on more robust and unambiguous methods for defining and "indicating academic words students need to know receptively, i.e. to enable them to read academic texts" (Pecorari et al., 2019, p. 61). As with the VLT-Ac, two equivalent forms of the AVT exist, both freely available at https://www.en.cityu.edu.hk/Vocabulary-Tests. The test format builds on the VLT-format (Nation, 1983, 1990; Schmitt, Schmitt, & Clapham, 2001) (Figure 12). However, the AVT consists of 19 multiple-choice clusters, equaling a maximum score of 57 points.

Figure 12

Example of an AVT-item (Pecorari et al., 2019)

_____ continue to do something	a. attain b. diminish
_____ keep something on its own	c. exploit d. induce
_____ reach a goal or objective	e. isolate f. persist

Note. Item example from the AVT by Pecorari et al. (2019).

The words in the AVT follow the frequency order of the AVL, which means that items at the beginning of the test are more frequent in academic texts than the items towards the end. This, in turn, makes it likely that test takers find the earlier words in the test easier than those towards the end.

In their validation of the AVT, Pecorari et al. (2019) administered the test to 455 university students "enrolled in undergraduate English studies at Swedish universities. The majority had Swedish as L1; a small number (less than 1% of the total) were exchange students … . Of the remainder, a minority were

bilingual" (2019, p. 63). In this case, the *majority* can range from 51 percent to 98 percent, and the *remainder* of bilinguals could be anything between 1 percent and 48 percent. More certain is the information about the students' choice of study discipline (English studies), indicating a strong possibility that the test takers' motivation to learn English and their already established English proficiency were rather high. Their university students had a proportional mean score of $M = .73$ on Form 1 and $M = .72$ on Form 2, causing Pecorari et al. (2019) to conclude that the AVT was fairly easy for their population. However, the data varied over a large span of test scores, which "suggest that the test has the potential to be useful with different, or less homogenous, populations" (Pecorari et al., 2019, p. 64).

Pecorari et al. (2019) calculated the internal consistency and got a Cronbach's alpha of .91 on Form 1 as well as on Form 2. Their results also confirmed the frequency distribution of the test with "fewer correct answers as frequency decreases" (Pecorari et al., 2019, p. 64). This should not be interpreted as a validity threat, but instead as fully expected, "since exposure to a word is necessary for learning" (Pecorari et al., 2019, p. 65) and frequency is the logic behind such an assumption.

Both the VLT and the AVT are tests of meaning-recognition knowledge and are said to tap into measuring partial lexical knowledge. In their revision of the VLT, Schmitt et al. (2001, p. 59) defined seven principles to be considered for the construction of the test items. In their third principle, the issue of partial knowledge and how it relates to the construction of a cluster was presented:

> Words are learned incrementally, and tests should aim to tap into partial lexical knowledge … . The Levels Test was designed to do this. The option words in each cluster are chosen so that they have very different meanings. Thus, even if learners have only a minimal impression of a target word's meaning, they should be able to make the correct match. (p. 59)

When Pecorari et al. developed the AVT it was "a process which was heavily informed by Schmitt et al.'s" (2019, p. 61). Therefore, the partial knowledge aspect was maintained as a construction principle also for the AVT:

> Differentiation among words in a cluster was a third principle. Two words in a cluster having similar meanings could lead to ambiguity. It was therefore necessary to keep words with closely related meanings or similar definitions (e.g., paraphrase and formulate) from occurring in the same cluster. When random selection put such words in the same cluster, one was either moved to another cluster or discarded. (p. 62)

A number of suggested score thresholds have been proposed in connection with the VLT. A threshold is believed to indicate mastery of the underlying word register for each section and "mastery of a 1,000-word band is determined by correctly answering a stated percentage of the items in a band" (McLean, 2021, p. 127). For instance, one commonly used mastery score threshold for the VLT is 26/30, indicating 86.67 percent correct answers, which is a percentage suggested in Schmitt et al.'s revised VLT (2001). Webb et al. (2017) proposed a stricter threshold set at 29/30, or 96.67 percent, for the high-frequency sections in their updated VLT. Since the VLT tests a basic aspect of word knowledge, aiming at "simply recognizing their form-meaning connections", Webb et al. argued that it may "be better to … use a higher cutting point for mastery" (2017, p. 56). Other mastery thresholds have been suggested (e.g., Nation, 1983; Olmos, 2009), and although there seem to be arbitrary uses of these thresholds, the 86.67 percent and 96.67 percent thresholds are those most commonly encountered in the literature. However, since testing always depends on the purpose and the context, different thresholds may be argued for different frequency bands (Webb et al., 2017). Although the VLT-Ac is not based on a frequency band in the same sense as the other sections, the above-mentioned thresholds have been used in studies focusing on scores on the VLT-Ac (e.g., Coxhead & Boutorwick, 2018; Durrant, 2016; Edgarsson, 2018; Skjelde & Coxhead, 2020). The developers of the AVT did not present any means of interpreting an AVT score in terms of mastery of the underlying AVL or how a score on the AVT – being a new instrument in the academic vocabulary testing field – could be compared to a score on the existing and widely-used VLT-Ac.

Both Schmitt et al. (2001) and Pecorari et al. (2019) followed up the administration adding interview data with a selection of participants to explore the behaviour of the VLT and the AVT in a stimulated-recall procedure. During this testing procedure, the participants were presented with target words from the written receptive meaning-recognition matching vocabulary test they had been administered. They were then asked to orally recall the meaning of the presented target word forms. One of the objectives of the follow-up session was to explore the extent to which students' correct responses on the meaning-recognition test depended on true word knowledge rather than other construct-irrelevant test strategies. In both studies, a conclusion was made that students who responded correctly on the written meaning-recognition test but who could not recall the meaning orally during the interview were plausibly guessing, thus inflating their scores on the written test, and not reflecting their true word knowledge (Schmitt et al., 2001, p. 76; Pecorari et al., 2019, p. 65). Although some of the Schmitt et al. (2001) interviewees reported having guessed, criticism could be raised towards the adopted procedure. As mentioned in passing in section 3.2.2., assessing meaning-recognition in a written test could be considered a different construct from that of assessing orally recalled word meaning (Laufer & Aviad-Levitzky, 2017; Laufer & Goldstein, 2004; Read,

2000). One evident argument for this criticism is that meaning-recognition "requires a lower threshold of lexical knowledge than meaning-recall" (Stoeckel et al., 2021, p. 188). With reference to Nation's table in section 3.1.2, a word knowledge taxonomy can be discerned, for instance, that recalling a word form's meaning productively is located higher up in the word knowledge taxonomy than is recognising a word form's meaning receptively. Consequently, in relation to this word knowledge taxonomy, a test taker who barely but truly knows a word in a meaning-recognition test is not likely to be able to recall its meaning in a productive test.

In addition to the guessing possibilities mentioned above, there are other issues of criticism concerning the matching format of the VLT, which also could be brought against the AVT as it builds on the VLT-format. This test format has been criticised for reasons concerned with item dependency and the relationship between distractors and target words (Kamimoto, 2008; McLean et al., 2015; Webb, 2008). Furthermore, it is hard to know whether it is the single words, the definitions or the two parts together that constitute the test item (e.g., Read, 2000; Warnby et al., 2022). Despite these restrictions, the test format is seen as an effective tool for broadly measuring the aspect of word knowledge (Pecorari et al., 2019; Read, 2000).

5.2. Estimates of academic vocabulary knowledge

The following section reviews studies that have estimated upper secondary school students' written receptive academic vocabulary knowledge in contingent Nordic contexts.

In short, the overall picture in the studies reviewed is that, on average, upper secondary school students in Denmark (Henriksen & Danelund, 2015), Iceland (Edgarsson, 2018), Norway (Skjelde & Coxhead, 2020) and Sweden (Gyllstad, 2007; Sylvén & Ohlander, 2019) do not reach the lower suggested mastery threshold of the VLT-Ac, indicating that they have a level of academic vocabulary knowledge judged to be below the required level for adequate reading comprehension.

In the case of Sweden, the reviewed studies used relatively small samples, did not have academic vocabulary as the focus of inquiry and the data were collected between ten and twenty years ago (the beginning of 2000's for Gyllstad, 2007 and beginning of 2010's for Sylvén & Ohlander, 2019). For this reason, the present thesis adds contemporary, richer and more specific information about this area of Swedish upper secondary school students' knowledge of academic vocabulary.

In Denmark, Henriksen & Danelund (2015) administered the two versions of the full VLT, that is: 2K, 3K, 5K, 10K and VLT-Ac (Schmitt et al., 2001),

to two samples of upper secondary school students, one on entry to upper secondary school (N = 26) and one after the final mandatory English course (N = 29). The results on the VLT-Ac are of particular interest in this thesis. Henriksen and Danelund (2015) used the lower 86.67 percent mastery threshold in their analyses. For their first (entry) sample, the mean proportional VLT-Ac score was M = .39 (SD = .28) and 92 percent of these students did not reach the 86.67 percent threshold. For their second (exit) sample, the mean VLT-Ac score was not reported, however, 69 percent of the students did not reach the lower 86.67 percent mastery threshold. Furthermore, as they included all sections, they observed higher vocabulary scores for all sections in the exit-sample than in the entry-sample. Although they concluded that the vocabulary development across grade levels was a positive result, the "fairly low level" (p. 49) of receptive vocabulary knowledge "of the learners is somewhat discouraging, considering the number of hours of language teaching they have had" (p. 41). In trying to explain this low level of vocabulary knowledge, Henriksen and Danelund point to the Danish educational system which emphasises oral skills and "stresses the use of communicative strategies to overcome language problems. In such a communicatively oriented, meaning-based learning environment, it is possible that learners may not have been encouraged to develop their vocabulary" (p. 50). The very large number of students not reaching the mastery threshold after upper secondary English instruction is a cause for concern for Henriksen and Danelund since many "of these students will go straight on to tertiary education, where many of the textbooks are in English … that require a vocabulary of considerable size … for the students to be able to cope with the academic demands" (p. 52).

In Iceland, Edgarsson (2018) administered the VLT-Ac to two samples of upper secondary school students after their final mandatory English course. For the combined sample (N = 417), Edgarsson observed a proportional mean test score of M = .84 (SD = .14) which indicates a possible ceiling effect and, more interestingly, that the average student did not reach the lower 86.67 percent mastery threshold. Citing, for example, the 88.8 percent threshold in Read (2000) or the 90 percent threshold in Tschirner (2004) and Olmos (2009), Edgarsson discussed the choice of threshold before selecting the 86.67 percent threshold in his study "as the minimum score for the participants to be judged as having acquired a large enough receptive knowledge of academic words in English for basic comprehension" (2018, p. 106). Edgarsson claimed that, although the .84 mean result may seem "quite impressive (. …), this result is not very encouraging" (2018, p. 106) and that "it appears that the average student in secondary school (from 18 to 20 years old) in Iceland has not yet acquired adequate knowledge for good understanding of the kind of vocabulary that is prevalent in academic language" (2018, p. 107). Moreover, Edgarsson also invited one of the sample groups (n = 168) to take an academic reading task (IELTS), the results of which revealed the same pattern (see 5.3.1. below).

Edgarsson concluded that despite "the tremendous amount of English language instruction the students have enjoyed throughout primary and secondary school" (2018, p. 109) there were large numbers of students scoring below the lower 86.67 percent VLT-Ac threshold and below a reading cut-off score. For this reason, Edgarsson argued for more research in how English academic literacy skills can efficiently be developed "under formal instruction in the education system for example through more content based instruction" so that the students may gain "the proficiency in academic vocabulary and discourse for success at the tertiary level" (2018, p. 111).

In Norway, Skjelde and Coxhead (2020) administered the two versions of the VLT-Ac to 134 first-year upper secondary school students "taking college preparatory courses, including their final obligatory English course, before qualifying for university studies" (p. 9/20). In addition to mean scores of the combined test versions (i.e., max 60 points), Skjelde & Coxhead also investigated the outcomes according to the lower 86.67 percent as well as to the higher 96.67 percent mastery threshold. As indicated by the mean, standard deviation and the median reported, it can be concluded that there was a ceiling effect in the scores. Skjelde and Coxhead observed a proportional mean VLT-Ac score of M = .81 (SD = .14) among their test takers where 58 percent did not reach the lower threshold and only 8 percent reached the stricter 96.67 percent threshold[19]. Their results lend evidence to previous Norwegian research where a lack of vocabulary knowledge was stated as a hindrance to students' academic reading proficiency (e.g., Busby, 2020a; Hellekjær, 2005, 2009). In line with Edgarsson (2018), Skjelde and Coxhead acknowledged the seemingly high level of academic vocabulary knowledge (.81) but noted that the variability in the data was large and that the number of students scoring below the lower mastery threshold (58%) indicated "a lack of appropriate receptive knowledge for these learners that can lead to difficulties with reading" (p. 14/20).

In Sweden, Gyllstad (2007) and Sylvén & Ohlander (2019) used the VLT-Ac among upper secondary school students. In a cross-sectional design, Gyllstad administered the VLT to two classes, one in the first year (N = 26, 16 years) and one in the second year (N = 28, 17 years) of upper secondary schooling. One of the conclusions drawn was that, in terms of mean difficulty on each VLT section, the VLT-Ac fits "neatly between the 3K and the 5K levels" (p. 140) affirming the placement of academic vocabulary along the mid-frequency vocabulary register. As to the students' academic vocabulary knowledge, the first-year students had a proportional mean score of M = .70

[19] Through personal communication, Skjelde has provided the mean and standard deviation of Version 1 of the VLT-Ac. The figures are the same, M = .81 (SD = .14). Furthermore, as indicated by the results but not explicitly expressed in Skjelde and Coxhead (2020), Skjelde acknowledged that the score distribution was negatively skewed, that is, indicating a clear ceiling effect within their sample.

(*SD* = .16) whereas the second-year students had a proportional mean score of *M* = .61 (*SD* = .16), which was significantly lower ($p < .05$) than the first-year group. Moreover, Gyllstad concluded that both upper secondary school student groups had problems with the words in the academic section as well as with the sections above 2K, and if "assuming that mastery of a level presupposes a score of at least 25 out of 30 [i.e., 83.33%], then these groups only reached mastery of the 2000-word-level" (p. 150). Gyllstad did not hypothesise about these outcomes or try to explain them, probably since his thesis focus was to develop a collocation test. Additionally, the significantly lower result in the second-year group possibly indicates a decrease in academic vocabulary development, however, the sample sizes are relatively small for drawing any stronger conclusions about development. In a longitudinal design, however, Sylvén & Ohlander (2019) reported a decrease in academic vocabulary knowledge during Swedish upper secondary schooling.

Sylvén & Ohlander (2014, 2019) explored upper secondary school students' English receptive vocabulary knowledge in a research project collecting data over three years. A first study (2014) presented the results from the students' first year but the final results were reported as part of the whole project in Sylvén (2019). In Sylvén and Ohlander (2019) the focus of inquiry was to compare English language development during upper secondary education in Content-and-Language-Integrated-Learning (CLIL) classes versus non-CLIL classes including gender analyses[20]. In the regular non-CLIL classes (reflecting the sampling for this thesis), Sylvén and Ohlander (2019) administered the VLT to two classes (N = 52) at two time points, firstly, on entry to upper secondary school (entry-sample) and secondly on completion of upper secondary education (exit-sample). The entry-sample had a proportional mean VLT-Ac score of *M* = .60 (*SD* = .18) and the exit-sample *M* = .57 (*SD* = .30). In the overall VLT score, there was no difference in mean scores between the two time points indicating that "the non-CLIL students' receptive mastery of L2 English vocabulary [had] stagnated" (p. 109–110). Furthermore, they observed higher VLT-Ac (as well as overall VLT) scores among males than females at both time points, but the difference was only significant for the VLT-Ac at the first time point ($p < .01$). Finally, with respect to the 86.67 percent threshold, it can be concluded from their data that the average final-year non-CLIL student performed below this lower threshold on the VLT-Ac. It must be noted that the main interest of investigation in Sylvén and Ohlander (2019) was to compare VLT-scores between students in immersion programmes (so-called CLIL-students) and students in non-immersion programmes. Interest-

[20] It remains unclear how balanced the English medium-of-instruction is between the CLIL-classes. In Sylvén and Ohlander (2014) it is stated that the CLIL programmes "where English, apart from being a separate subject, is also the medium of instruction in several or all subjects, e.g. biology and history" (p. 84).

ingly, even the CLIL-students' average VLT-Ac score in their final upper secondary school year was below the lower mastery level (*M* = .83, *SD* = .07, *N* = 57). For these reasons, Sylvén and Ohlander (2019) concluded that:

> the most important pedagogical conclusion to be drawn from the results presented in this chapter is that the teaching in non-language subjects where English is used as the medium of instruction, as well as in English as a subject of its own, should focus a great deal more on academic vocabulary proficiency. (p. 115)

To summarise, the research conducted in the Nordic countries suggests that upper secondary school students exhibit a range of abilities in academic vocabulary knowledge. Some students score high and some score low. On average, the Nordic upper secondary school students do not reach the suggested lower mastery threshold of 86.67 percent. However, the Swedish sample sizes in Gyllstad (2007) and Sylvén and Ohlander (2019), and the Danish (Henriksen & Danelund, 2015) and Norwegian (Skjelde & Coxhead, 2020) sample sizes are too small to draw any larger conclusions or generalise from. Furthermore, Gyllstad (2007) collected his data at the beginning of the millenium and Sylvén and Ohlander at the beginning of the 2010's. For this reason, it is worthwhile investigating Swedish adolescents' academic vocabulary knowledge with a larger sample approximately 10 to 20 years later, especially since a new national curriculum was implemented in 2011. It should be noted also that Gyllstad's aim (2007) was to develop collocation tests and Sylvén and Ohlander (2019) focused on the language gains from being a CLIL student (target group) in comparison to a non-CLIL student (control group). Thus, neither of these Swedish studies had academic vocabulary as the main object of investigation. To better understand the relevance of academic vocabulary it is also important to understand how this specific construct may relate to other factors. The following section reviews some of the knowledge we have about these relationships.

5.3. Factors related to academic vocabulary knowledge

Besides a focus on receptive academic vocabulary knowledge, research relating to students' academic literacy may involve investigations of, for example, their disciplinary literacy (Airey, 2009), their reading strategies (Busby, 2018) or their productive use and knowledge of academic vocabulary (Edgarsson, 2018; Olsson, 2016). Research on academic vocabulary knowledge may correlate vocabulary test scores with other proficiency variables, such as reading or achievement; demographic variables, such as L1 or gender; or behavioral variables, such as spare time reading or gaming. This section reviews studies using measures of receptive academic vocabulary in relation to such variables.

5.3.1. L2 academic vocabulary and reading proficiency

This thesis focuses on academic vocabulary as an essential facet of academic reading literacy. With respect to the 98 percent lexical coverage (see section 3.3.) and the fact that academic vocabulary constitutes approximately 10–14 percent of the words in academic texts (see section 4.2.), it can be assumed that academic word knowledge is essential for academic reading and that variation in academic vocabulary knowledge may explain a meaningful amount of the variation in reading proficiency. As empirical proof, correlations have been established between receptive academic vocabulary scores and academic reading comprehension scores which have shown to be positive. Shaw and McMillion (2011) correlated scores on a recognition test targeting AWL words with scores on an academic reading task among 80 Swedish first-year university biology students. They reported a statistically significant correlation of $r = .66$.

A similar value, $r = .70$, was observed by Edgarsson (2018) who used VLT-Ac scores with academic reading scores among 168 Icelandic upper secondary school students. Edgarsson administered two academic reading tasks from a freely available practice test from the International English Language Testing Service (IELTS). Under normal testing conditions, the reading booklet of the full IELTS test consists of three tasks (three texts + questions). Edgarsson defined the two chosen texts as being of a general nature and "characterized by academic language both in terms of sentence structure and vocabulary" with a "high ratio of AWL words" (2018, p. 103). IELTS scores are divided into proficiency bands ranging from the lowest band 1 to the highest band 9. Across the globe, universities that welcome English L2 students often demand a minimum English proficiency of IELTS band 6 (or equivalent). Edgarsson (2018) observed a mean band score of 5.6. However, to better understand how many students reached band 6, he categorised his participants into the nine IELTS proficiency bands and found that 40 percent of the upper secondary school students did not reach IELTS band 6. He paralleled this finding with the findings of Hellekjær (2009) who found that 33 percent of his Norwegian upper secondary school students scored below band 6. Furthermore, it should be noted that the band 6 level may probably be too low a requirement for achievement in English medium-of-instruction university studies (e.g., Feast, 2002; IELTS, 2022; Trenkic & Warmington, 2019). A final remark (see section 5.2.) about the Edgarsson (2018) study is that, in addition to the observed 40 percent under-achieving students, about 50 percent of the students scored beneath the lower 86.67 percent mastery level of the VLT-Ac. This confirms the predictive value of academic vocabulary in relation to academic reading literacy.

A study related to these findings was conducted by Masrai (2019) where high-, mid-, and low-frequency vocabulary scores were correlated with com-

prehension scores on an academic IELTS reading test. Since the academic vocabulary can be seen as part of a mid-frequency measurement of overall vocabulary size (see section 3.1.2), it is not surprising that the highest correlation was found between mid-frequency vocabulary and reading comprehension ($r = .77$, $p < .01$). Although the mid-frequency test battery in this study included academic words (in AWL/AVL) and other mid-frequency words, the finding indicates the strong relationship between the mid-frequency vocabulary and academic reading. High-frequency vocabulary knowledge is always important, but by disentangling the predictive power of the three frequency levels Masrai (2019) was able to conclude that "both high- and mid-frequency vocabulary were observed to contribute uniquely to [English academic] L2 reading comprehension" whereas low-frequency vocabulary "did not appear to explain any additional value in the predictive model of L2 reading performance" (p. 10).

Since vocabulary and reading has a reciprocal relationship, it is not just vocabulary that explains reading comprehension, but reading also explains vocabulary development. As regards academic vocabulary and academic reading, this was investigated by Freimuth (2020) among 13 L2 English university students in a one-semester course in English for academic purposes. Freimuth argued that it is an "unreasonable expectation" (p. 32) that all the AWL words be taught during one semester, and, for this reason, wanted to investigate whether exposure to a variety of academic texts would increase students' academic vocabulary knowledge incidentally. From the texts (policy papers, book chapters, essays, journal articles ranging between 750–7000 words) AWL words "were chosen for their level of difficulty and common usage in academic writing as determined by the researcher" (p. 36). A total of 13 words were used and assessed according to an adaptation of the VKS (Paribakht & Wesche, 1993):

> I. I have never seen this word before.
> II. I have seen this word before but don't know what it means.
> III. I have seen this word before and I think it means _____.
> IV. I know this word. Here is a sentence with it: _____ __.
> (Freimuth, 2020, p. 36)

From pre- to post-test, the findings revealed, among all but one student, an increase in academic vocabulary knowledge as a by-product of meaning-focused academic reading. Freimuth concluded that incidental academic vocabulary learning indeed takes place through reading academic texts "despite the different frequencies of exposure to the word, which ranged from 1 to 4, a relatively low exposure rate" (2020, p. 39).

The reviewed studies all point to the essential role that academic vocabulary knowledge plays in academic reading success. In addition to general high-

frequency vocabulary knowledge, the unique contribution academic vocabulary knowledge makes to academic reading comprehension validates the use of academic vocabulary measures for predicting students' ability to engage with academic reading tasks. Furthermore, there is empirical evidence that academic vocabulary can be incidentally learned through academic reading.

5.3.2. L2 academic vocabulary and academic achievement

Since an overall large vocabulary size is central to good language proficiency, and language proficiency, in turn, is central to academic achievement (Masrai & Milton, 2017), the role that academic vocabulary knowledge may play for achievement is fundamental for this thesis. Masrai and Milton (2018) administered an AWL test and a test of overall vocabulary size to a sample of Arabic-speaking EFL students to investigate the relationship vocabulary knowledge may have with academic performance. As a measure of academic performance, they used student grade point average (GPA). They found that both vocabulary tests explained the variance in GPA. However, the strong, positive correlation between participants' academic performance and receptive academic word knowledge ($r = .73$, $p < .001$) indicated that "AWL words do fulfil some important function in addition to maintaining a vocabulary of the right general size, which allows the possessors of this vocabulary to gain better GPAs" (p. 54).

Similar patterns have been found among lower secondary language minority English L2 students (Townsend et al., 2012) where scores on the VLT-Ac were used together with an aggregated score on a test of basic skills in reading, mathematics, civics and science. Although the explanatory power of academic vocabulary was lower than in Masrai and Milton (2018), Townsend et al. (2012) concluded that since academic word knowledge explained "unique variance in academic achievement (. …) this finding lends credibility to the construct of general academic word knowledge and its importance for academic success" (p. 513).

In parallel, Skjelde and Coxhead (2020) correlated Norwegian upper secondary school students' VLT-Ac scores with their English grades and found a moderate, but statistically significant and positive correlation of ($\rho = .37$, $p < .001$). Given that they observed a clear ceiling effect in the VLT-Ac scores which may attenuate the correlation, the explanation is still meaningful and Skjelde and Coxhead argued that this finding indicated that "academic achievement correlates with academic vocabulary knowledge for L2 English learners" (p. 15/20). Moreover, they found that students with higher grades had a significantly higher probability of reaching mastery thresholds of the VLT-Ac.

Although these three studies used different groups of participants and not exactly the same measures, they all provide evidence that academic vocabulary knowledge offers a logical and convincing explanation for academic

achievement and, hence, is important for learners to acquire in their development of academic English.

5.3.3. L2 academic vocabulary and extramural English

Quite a lot is known about the relationship between general vocabulary knowledge and exposure from extramural English (EE), but the research field has not offered much information on the relationship between academic vocabulary knowledge and extramural English involvement. A review of the literature suggests that there are only two studies specifically focusing on extramural (or extracurricular/out-of-school and/or leisure/spare-time) English and academic vocabulary (Busby, 2020b; Olsson & Sylvén, 2015).

In the Swedish context, Olsson and Sylvén (2015) explored the impact that extramural English habits could have on the productive use of academic vocabulary. They analysed upper secondary school students' writing assignments at two time points by counting the number of academic words in the students' texts. It was found that males had a higher number of academic words in their texts than their female peers at both time points and that males were involved with extramural English activities to a higher degree than females. However, the growth curve difference between genders was not significant. For this reason, Olsson and Sylvén concluded that EE did not affect the productive use of academic vocabulary. However, their results indicated that exposure to books, to computer games and/or to movies may affect students' written productive use of general lexis.

In Norway, Busby (2020b) administered a digitalised version of the full VLT with its five sections together with an EE survey to students in their first, second, third, or fourth year of university studies. The EE survey collected information about the students' frequency of reading books, reading on-line and gaming. All three EE factors affected scores on the full VLT, however, when separating the VLT-sections she found no statistical significance that any EE factors may predict academic vocabulary knowledge. It must be noted, though, that Busby observed a high ceiling effect with the VLT-Ac in her sample of university students which may have affected the analytical outcomes.

With only these two studies, our knowledge about the relationship between academic vocabulary and EE is very limited and previous scholars call for more research about how EE can be a source for different vocabulary categories, for example, academic vocabulary (Puimège & Peters, 2019; Schmitt, 2019). Olsson and Sylvén (2015) did not find any significant relationship between EE and students' own productive writing. Busby (2020b), on the other hand, explored EE with written receptive academic word knowledge. However, due to ceiling effects, too little variation in the VLT-Ac scores was likely to have affected the explanatory power of EE exposure, and for this reason, Busby (2020b) suggested that the AVT may be a more appropriate measure

of students' academic vocabulary knowledge. Thus, prior to this thesis, it can be concluded that no studies have investigated the relationship between different EE activities and written receptive academic vocabulary knowledge among upper secondary school students.

5.4. Chapter summary

This chapter has reviewed research that has developed testing instruments targeting written receptive academic vocabulary knowledge. Two tests – the VLT-Ac and the AVT – were given most attention. The review continued by presenting results from studies using the VLT-Ac among upper secondary school students in the Nordic context – relevant for the purposes of this thesis. The overall picture suggests a low level of academic vocabulary knowledge. However, the sample sizes in the reviewed studies conducted in Sweden, Denmark and Norway may be too small to generalise from; the results from Iceland seem more robust considering the larger sample size. Moreover, previous studies investigating the relationship that academic vocabulary may have with reading and with achievement show a positive correlation which strengthen the validity argument of using academic vocabulary knowledge measures as an important factor of academic reading comprehension. However, studies focusing on the association that extramural English may have with academic vocabulary show no relationship. Significantly though, in one of the studies, ceiling effects in the VLT-Ac scores may have attenuated the correlation. In general, there is little research about academic vocabulary knowledge in this population, in relation to achievement and to extramural English.

6. Method

This chapter describes the research design adopted in the various studies included. First, the overall research design is presented in section 6.1. In section 6.2., methodological issues linked to the choice and recruitment of participants are discussed. Section 6.3. introduces the instruments used: the English Academic Vocabulary Test (AVT), the academic section of the Vocabulary Levels Test (VLT-Ac), the Background questionnaire, the Extramural English survey, and Register data. Section 6.4. presents findings from the pilot phases. Section 6.5. discusses the data analyses used in the three studies. The last two sections address issues regarding validity (6.7) and ethics (6.8).

6.1. Research design

The thesis adopts a cross-sectional design and uses statistical analyses on data from two different sample groups connected to the overarching research topic of investigating EFL learners' written receptive English academic vocabulary knowledge. To answer the two research questions about what the students' academic vocabulary knowledge is and what explains this knowledge, three aims were expressed (section 1.2.). To measure these students' academic vocabulary knowledge, the first aim was to identify an appropriate measurement instrument for the target population. Having identified a suitable instrument, a second aim was to map the knowledge and development of students' academic vocabulary knowledge by employing measurement data from two time periods: 1) at the beginning of upper secondary education at the time of the first mandatory course, English 5 (1st time period = TP1, *entry-sample*), and 2) after the last mandatory course, English 6, at the beginning of year 3 (2nd time period = TP2, *exit-sample*). A third aim was to investigate what factors may explain students' academic vocabulary knowledge. At both time periods, data were collected using paper-and-pencil administered vocabulary tests and questionnaires that were later manually processed into digital data sets used for statistical analyses.

The principal purpose of collecting repeated cross-sectional data was to estimate academic vocabulary development during mandatory upper secondary English instruction. A longitudinal design would undoubtedly have made a stronger case, but, mainly due to time restrictions, a cross-sectional design

was adopted. Below, Table 5 presents the data collection time periods (TPs) with details of sample sizes and instruments used in the thesis.

Table 5

The two time-periods of the cross-sectional data collection

	TP1[1]	**TP2**
Sample	Entry-sample	Exit-sample
Born	~ 2003	~ 2002
Time	Aug – Sep 2019 Beginning mandatory EFL instruction.	Aug – Sep 2020 Exited mandatory EFL instruction.
N	552	446
Instruments	AVT Background questionnaire 1 EE survey	AVT Background questionnaire 2 & 3 EE survey VLT-Ac

[1] TP = Time period

6.2. Data collection procedure and sample

An intention set for the data collection was to gather data from each programme represented at least in two schools with a minimum number of around 400 students in total per sample. This sample size is in line with the size selected as representative for the target population within the construction of the national assessment exams in English provided by the Swedish NAE (see, for example, Erickson et al., 2022). During Spring 2019, twelve schools were sent an invitation letter (Appendix A).

One requirement for participation was that each school promised to provide N classes from N programmes at both time periods, for example, if a school participated with one Social sciences (SA) and one Natural sciences (NA) class in TP1, the parallel classes for TP2 should be included. Within these constraints, seven schools chose to participate. Within the sample, there are schools from large cities as well as small cities. Both municipal and independent schools are included.

All students were informed about the research in advance through a letter sent out to the teacher/school principal. The information and consent letter were read aloud by the researcher to the class on the day of data collection and students were given the opportunity to ask questions. All students were above 15 years of age and, therefore, were able to consent to participate themselves

by signing individual consent letters (SFS 2003:460). An example of a consent letter is appended (Appendix B).

For the Entry-sample, the data collection (TP1) took place in Autumn 2019. A total of 22 first-year classes were visited once during their 3rd to 9th study week. The time allotted for data collection in each class was set to 60 minutes excluding 5-10 minutes of instruction. All classes were administered the AVT. Moreover, the following tasks were included:

- a background questionnaire (L1, gender, parental educational level, …),
- a self-estimation survey of current extramural target language exposure.

Together with the teachers, it was decided that all students stayed in the classroom and took part in the data collection process. This meant that students who were unwilling to participate or did not consent to use their booklets most often remained in the classroom. Their booklets were handed in to the researcher and destroyed. A total of 552 students consented.

For the exit-sample, that is: the total of programmes, classes and students which had contributed to the entry-sample, the data collection took another direction than was primarily intended. Originally, TP2 was planned for Spring 2020, however, at that time schools were in lock down because of the Covid 19 pandemic. A new design was therefore devised. Rather than collecting data at the end of the last mandatory course, the new exit-sample was comprised of students who had *exited* the last mandatory English course, English 6, and thus had just completed their second year of upper secondary schooling. This means that these students were visited directly at the beginning of their third year, which was approximately 9 weeks after completing English 6 in June 2020. This was a summer when students were required to remain in Sweden, due to Covid 19 travel restrictions, and were not able to improve their English language by making trips abroad. The same seven schools as in TP1 were visited, however, with 20 classes instead of 22 since two classes were taught from their homes on the planned date. Therefore, two additional classes were recruited externally. The data collection took place during the 2nd to the 8th week in year 3. According to the number of students given by the teacher in advance, approximately 550 students should have been visited. However, due to a further period of Covid restrictions, the number decreased, on average by five students per class. A total of 446 students consented and took part in the data collection. Furthermore, the time available in each class varied. For some classes, 40 minutes were allotted, and for others 60 minutes. For the two additional classes, only 30 minutes were given. This also changed the choice of

instruments to be used. All classes took the AVT. In addition, the following instruments were included:

- the VLT-Ac,
- a background questionnaire (short or long version)
- a self-estimation survey of current extramural English exposure

A total of 998 students consented. For each study, a selection from this full dataset of 998 participants was made depending on the inquiry of the study. For instance, Study II used data from 952 students and Table 6 presents the distribution of those sample participants in relation to the national population distribution (N ≈ 200,000 students in university-preparatory programmes; Skolverket, 2022d).

Table 6

Distribution of participants in relation to the distribution in the national population (Study II)

		ES	HU	NA	TE	EK	SA
Total	Sample (*N* = 952)	171	68	137	103	171	302
	Prop. of tot. sample	.18	.07	.14	.11	.18	.32
	Prop. in the nat. target pop. (*N* ≈ 200,000)	.10	.01	.21	.14	.23	.29
Share of females	Prop. of sample prog.	.80	.85	.55	.27	.39	.60
	Prop. in the nat. prog. pop.	.63	.79	.55	.19	.51	.65
Share of tertiary educated parents	Prop. of sample prog.	.67	.79	.70	.74	.74	.56
	Prop. in the nat. prog. pop.	.67	.69	.77	.72	.65	.63
Sample groups	Entry (*n* = 526)	102	36	73	56	110	149
	Exit (*n* = 426)	69	32	64	47	61	153

Note. Prop. = Proportion; tot. = total; prog. = programme; nat. = national; pop. = population.

6.3. Instruments

The following section describes the instruments used in the three studies.

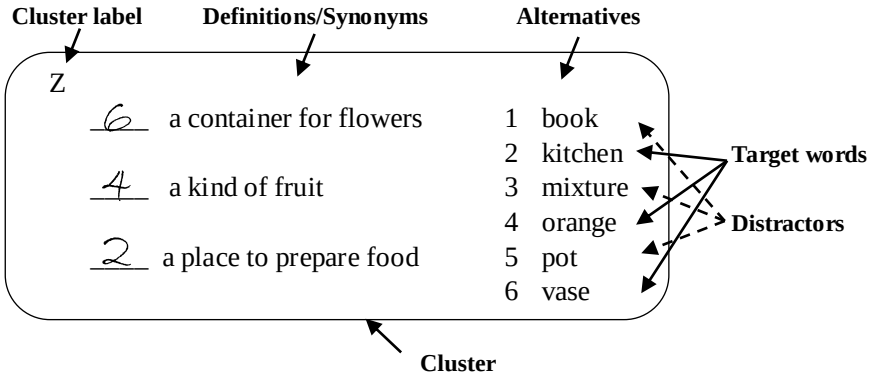
6.3.1. The Academic Vocabulary Test – AVT

Form 2 of the Academic Vocabulary Test (AVT) (Pecorari et al., 2019), was used as a measure of written receptive English academic word knowledge in all three studies. The AVT targets 57 academic words which also is the maximum possible test score (See Chapter 4.2. and 5.1.).

Figure 13 illustrates how the AVT-format was used in this study, how the test is completed and how its different components are defined.

Figure 13

Example of the AVT-format used in this study, together with completed answer and illustrating the different components.



As was noted in section 5.1., the AVT (as well as the VLT-Ac) is a test favoring partial lexical knowledge and, therefore, the construction of an item cluster should differentiate clearly between the six alternative words, thus, preventing words with "closely related meanings or similar definitions … from occurring in the same cluster" (Pecorari et al., 2019, p. 62). However, during the exploratory IRT analyses (2PL and 3PL models) in Study I (Warnby et al., 2023), two items in Form 2 of the AVT were found to violate this principle, namely item 44 (cluster O) and item 49 (cluster Q) (Figure 14). For these two items, mean score comparisons were made, and lexical analyses were performed using the Merriam-Webster and the MacMillan online dictionaries.

Figure 14

The AVT clusters O and Q revealing two outliers (item 44 and 49)

O

_____ just starting$_{43}$	1	ancillary
	2	asymmetrical
_____ providing support$_{44}$	3	defensible
	4	feudal
_____ travelling from place to place$_{45}$	5	incipient
	6	itinerant

Q

_____ attach$_{49}$	1	append
	2	catalyze
_____ make something start to happen$_{50}$	3	delimit
	4	interconnect
_____ mark the boundaries of something$_{51}$	5	oversimplify
	6	reformulate

According to the key provided by Pecorari et al. (2019, https://www.en.cityu.edu.hk/Vocabulary-Tests), the definition for item 44, *providing support*, should be matched with the option *ancillary*. However, in Study I, this item was an extreme outlier. For this reason, a mean score comparison was made for this cluster which revealed that 47 students with a group test mean score of M = 38.62 chose the target word *ancillary*. However, 127 students with a group test mean score (M = 31.76) above the total test mean score (M = 30.38) chose the option *defensible* (Table 7).

Table 7

Mean score comparisons for AVT item 44 definition "providing support": students choosing alternative 1 or 3.

Alternatives	***N***	**Mean AVT score**	***SD***
1 *ancillary*	47	38.62	10.991
3 *defensible*	127	31.76	10.338

A lexical analysis of the words in the cluster was therefore made using the Merriam-Webster online thesaurus which claims *defensible* to be synonymous

to *supportable* which is closely related to the given item definition *providing support* in the AVT:

> 2 capable of being defended with good reasoning against verbal attack // there's simply no *defensible* reason for dropping out of school
>
> **Synonyms for *defensible***
> defendable, justifiable, maintainable, supportable, sustainable, tenable (https://www.merriam-webster.com/thesaurus/defensible, retrieved 13 July 2022)

Reformulating the Merriam-Webster example sentence may be used as an argument that *defensible* is closely related to the definition *providing support*:

> there's simply no reason *providing support* for dropping out of school

Because of the partial knowledge argument, the students who chose *defensible* were also credited with one point.

The same pattern was observed for the second outlier item. The definition *attach* is, according to the original key, to be matched correctly with the option *append*. The test taker group, 77 students, who chose *append* had a mean score (M = 31.48) slightly above the total mean score (M = 30.38). However, the strongest group of test takers (M = 36.04) was a group of 186 students who instead chose the option *interconnect* (Table 8).

Table 8

Mean score comparisons for AVT item 49 definition "attach": students choosing alternative 1 or 4.

Alternatives	*N*	Mean AVT score	*SD*
1 *append*	77	31.48	11.291
4 *interconnect*	186	36.04	9.344

With reference to both Merriam-Webster and MacMillan neither 'append' nor *interconnect* are said to be synonyms to *attach*. However, when searching for *append* in Merriam-Webster, *attach* occurs as a related word, however, not in MacMillan. More interesting is the fact that, in MacMillan, one of the top ten synonyms for *interconnect* is the word *attach*, which is the given definition/synonym in item 49. For this reason, the 186 students who chose *interconnect* to be matched with *attach* were also credited with one point.

The scoring approach adopted for these two items was therefore defined as a *lenient* scoring approach in Study I, and the one in the original key (Pecorari et al., 2019, https://www.en.cityu.edu.hk/Vocabulary-Tests) was defined as a *severe* scoring approach. Future recommended revisions of the AVT might include a revision and/or a rewrite of these items. The lenient approach was used in Study I and II, whereas Study III (chronologically conducted before Study I and II) adopted the original severe scoring approach.

6.3.2. The Vocabulary Levels Test-Academic section – VLT-Ac

Version 1 of the Vocabulary Levels Test-Academic section (VLT-Ac) (Schmitt et al., 2001) was used in the pilot phase and in Study I. The VLT-Ac is similar to the AVT in test format in the sense that it targets academic vocabulary. However, the VLT-Ac has 30 academic target words sampled from the AWL (see sections 4.2. and 5.1.)

Since the VLT-Ac is a widely used test of written receptive English academic vocabulary knowledge it was judged necessary to link scores from the VLT-Ac with scores from the AVT to be able to make valid comparisons between studies. This was done in Study I. Figure 15 shows how the format of the VLT-Ac was used in this study.

Figure 15

Example of the VLT-Ac format used in this study

_____ change	1 achieve 2 conceive
_____ connect together	3 grant 4 link
_____ finish successfully	5 modify 6 offset

6.3.3. Background data and questionnaire

The background data build on register data provided by the schools and on responses in a background questionnaire.

The background data consist of programme (ES, EK, HU, NA, SA, TE) and self-reported information collected in the questionnaires, including items such as: personal identification number, gender, L1/s/, the number of years of EFL instruction, years in Sweden, time spent abroad in an English-speaking country, parents' educational level, parents' L1s, self-assessed reading ability and fluency, among others. Grades from the school's register were obtained

from all municipal schools. However, this was not possible for all independent schools. At the end of the questionnaire, students could leave open comments.

All students filled in a background questionnaire, although different questionnaire formats were used depending on time-period and allotted time in class.

Students in the entry-sample had the same questionnaire (Questionnaire 1) since the time given for data gathering in each class was constant, 60 minutes. An excerpt of Questionnaire 1 can be found in Appendix C.

The background questionnaire for the exit-sample differed slightly from that used for the entry-sample and contained some additional questions, such as self-reported grades in English 5 and English 6. Furthermore, since the time slots given in the classrooms for the exit-sample differed, some of the classes were given a full version (Questionnaire 2a) whereas others filled in a shortened version (Questionnaire 2b). Questionnaire 2a can be seen in Appendix D.

6.3.4. Extramural English survey

In the Extramural English (EE) survey (Appendix E), students estimated their current extramural English activities on a 5-point Likert scale with ordinal data ranging from 1 = *never or almost never* to 5 = *more than 2 hours almost every day*. This scale is a modified version of the 5-point scale used in the student questionnaire focussing on spare time reading in the large-scale study PISA in 2018[21] (OECD, 2019). The EE questions derive from the European Survey of Language competences (SurveyLang) (European Commission, 2012) used for example by Peters (2018) (Appendix F).

During the pilot stages, the SurveyLang EE questions were modified internally and expanded in number, for example, the items on reading books and magazines were separated into three categories: reading fiction, reading factual literature and reading news. In a pilot study (Warnby, 2021), it was found that participants' (N = 62) self-reported rates of reading fictional and factual literature (M = 1.2), and newspapers (M = 1) were low, whereas reading in English on the Internet had the second highest value (M = 3.4). The pilot participants commented that they read both news and factual texts but not in printed form. When adjusting the question according to genre instead of medium, for example, reading fiction regardless of whether it appears in paper books or e-books, reading news and articles printed or digital, the average value for each factor increased immediately. For this reason, the SurveyLang EE question *How often do you visit websites written in English* was omitted. Instead, the Internet was included as a source for watching TV and movies or

[21] PISA = Program for International Student Assessment. An excerpt from the questionnaire is presented in Appendix G. https://nces.ed.gov/surveys/pisa/pisa2018/questionnaires/Student_Q_Booklet_English.html

for reading fiction or faction (as in PISA 2018). The interest of this study is not whether EE takes place in front of broadcast TV and paper-printed text or in front of a digital device, but rather, whether the exposure to and interaction with English occurs outside school. Today, the mode of input from genres such as written fiction or motion pictures is easily reached digitally, especially since videos, movies, series and TV are most often streamed. Furthermore, the fact that they read a lot of English on the Internet every day was not particularly informative. Additionally, the final EE questions differentiated between audio-visual exposure supported by Swedish or English subtitles, and not just subtitles in general. In Sweden, almost nothing is dubbed (except programmes and movies for children). A total of fifteen EE factors were used in the analyses:

1. Watching movies and series with Swedish subtitles
2. Watching movies and series with English subtitles
3. Watching movies and series without subtitles
4. Watching factual programmes with Swedish subtitles
5. Watching factual programmes with English subtitles
6. Watching factual programmes without subtitles
7. Watching non-factual programmes with Swedish subtitles
8. Watching non-factual programmes with English subtitles
9. Watching non-factual programmes without subtitles
10. Reading fiction
11. Reading non-fiction
12. Reading news and newspapers
13. Listening to music with English lyrics
14. Listening to radio and podcasts
15. Gaming.

6.4. Pilot phases

During Fall 2018 and Spring 2019, three rounds of piloting for vocabulary tests and questionnaires were conducted. Four classes from three different upper secondary schools took part in these pilot stages. The vocabulary tests used were the academic section of version 1 of the revised VLT-Ac (Schmitt et al., 2001) and form 2 of the AVT (Pecorari et al., 2019).

The pilot rounds and results were presented at two conferences which gathered insights to improve the design before final data production (Warnby, 2019a, 2019b). Table 9 provides an overview of the three pilot rounds:

Table 9

The pilot rounds with details on programme, number of participants, month/year, versions and choices of vocabulary tests, and versions of questionnaire

Round	Prog.	N	M-Y	English test	Questionnaire
1	NA	19	Nov-18	VLT-Ac	Version 1
2	EK	25	Feb-19	AVT	Version 2
3	SA	17	Apr-19	AVT	Version 3
	SA	19	Apr-19	AVT	Version 3

In the first pilot round, one class (n = 19) took the 30-point VLT-Ac with results (M = 24.9, SD = 3.8; *proportional score* = .83, *proportional SD* = .12) indicating an imminent ceiling effect if the VLT-Ac was to be used in the main study where even older participants would be included. Similar ceiling effects have been seen when the VLT-Ac has been administered to, for example, Icelandic as well as Norwegian upper secondary school students (Edgarsson, 2018; Skjele & Coxhead, 2020) and to Norwegian university students (Busby, 2020a). In the two succeeding rounds, three classes (n = 59) took the 57-point AVT with clearly lower percentage results, however, with a large variability (M = 18.1, SD = 8.1; *proportional score* = .32, *proportional SD* = .14). The AVT was chosen as the main English test to be used, since it (*a*) appeared not to be affected by a ceiling effect, and (*b*) contained almost twice the number of items which strengthened the possibility of obtaining good internal consistency and variability. Completing the AVT took on average 10–14 minutes, whereas the VLT-Ac was quicker to finish, about 7–11 minutes.

The questionnaires aimed at collecting information from three main domains: (*i*) background biographical information, (*ii*) exposure to and interaction with some out-of-school language-driven activities, and (*iii*) reading ability and reading strategies. During all pilot rounds, students commented on irregularities, questions, instructions, and so forth., and expressed feelings regarding the test battery. These comments and reactions were discussed and considered, and frequently led to revisions of the questionnaires.

6.5. Data analyses

In all three studies, the data were analysed with the aid of statistical procedures including both descriptive and inferential statistics. Although SPSS was used in all three studies, the analyses in Study I mainly depended on estimations

performed in R Studio (RStudio Team, 2020). Additionally, the open comments students provided in the background questionnaire contributed to shedding light and adding information to the analyses. However, these data are not used as the main empirical data for the three studies.

In Study I, data from 385 students from TP2 were used to link scores on the two academic vocabulary tests (VLT-Ac and AVT). The linking procedure chosen was a concurrent calibration (Feuer et al., 1999; Kolen & Brennan, 2014) using Item Response Theory (IRT) modeling (Embretson & Reise, 2013; Hambleton & Swaminathan, 2013) performed in RStudio using the mirt package (Chalmers, 2012). In the preparatory analyses, Classical Test Theory (CTT) and factor analysis were used to assert the viability of conducting an IRT-based concurrent calibration. For these analyses, SPSS was used, which also was used, on occasion, for cross-checking.

In Study II and III, data from both time periods were analysed in SPSS. Descriptive statistics were presented, including measures of central tendency and dispersion. CTT-analyses included checking for internal consistency (Cronbach's alpha) and discrimination indices.

In Study II, a combined number of 952 participants was used to map the knowledge and the possible development of academic vocabulary knowledge during mandatory English instruction. In addition to the descriptive statistics, T-tests were used to investigate, for example, observed mean differences between different programmes. Furthermore, in a sub-sample (n = 413) of the students in TP2, correlational analyses were performed using scores on the AVT and final grades in the English 6 course. Finally, logistic binary regression analyses were used to investigate how different group variables may increase the probability of reaching suggested mastery thresholds as indexed by AVT-scores.

Study III focused on the frequency of extramural English activities among 817 students from TP1 and TP2 and how these different activities (EE factors) may be correlated with academic vocabulary knowledge. T-tests were used to investigate the significance of the observed differences in the amount of involvement with EE activities among, for example, males and females. Moreover, a linear regression model was used to explore the explanatory power that EE factors, gender, number of L1s, years of English instruction, and parental educational level may contribute to the variance in academic vocabulary knowledge.

Although the relatively large sample sizes and the choice of statistical methods were useful for addressing the purpose of the thesis, a critical perspective would display a number of issues regarding the validity and generalisability of the results. Some of these issues will be dealt with in the following section.

6.6. Validity issues

The data analyses in the three studies are heavily dependent on statistical procedures providing descriptive and inferential results that are qualitatively discussed in the light of their external validity formulated within the thesis' overriding research questions. Hence, some general comments regarding the possibilities and constraints of the different statistical analyses are worth mentioning.

Descriptive statistics are not only used to describe a data set but also to provide a basis for inferential analyses. Descriptions can display what has been observed but are unable to provide any explanation. The explanation has to be argued through further elaborations. Correlational and differential statistics can provide such arguments. The bigger the sample is, the stronger the explanatory arguments are for the target population. Correlational research focusing on the relationship between variables must therefore have precise variable measurement procedures. Differential research, having parallel principles as above, focuses explicitly on comparisons between sub-samples defined by a pre-existing variable; all groups need to be handled identically.

However, inevitably there are always errors of measurement in research. One dilemma of external validity in research concerns the number of participants. If the number is too small it can affect the generalisability, making subsample and categorical analyses difficult. In large-scale research, numbers from 10,000 and up are not unusual, but some regard a sample size of as few as 30 participants as enough (Field, 2013). Certainly, it all depends on the inquiry. In Sweden, the national test development for English is developed using a variety of quantitative and qualitative methods (Erickson, 2010; Erickson et al., 2022). For example, all tasks and items are pre-tested in "large, randomly selected groups, normally comprising around 400 students" (Erickson, 2010, p. 3) and group comparisons are made. Sample size also pertains to the number of items in a vocabulary test, which is an issue dealt with in Chapter 5.

Another typical dilemma for statistical analyses is when the data show little variability since this is a requirement when searching for powerful correlations. Data can be highly skewed causing possible problems with, for example, ceiling effects which may increase the number of Type I errors (Austin & Brunner, 2003). A pedagogical view on such a negatively skewed distribution of results on a proficiency test could be positive since the test-taking learners, in line with equity and educational objectives, all seem to be very proficient. However, such a distribution would fail to inform on other researched variables. For this reason, a research tool must be able to show individual differences, but long tests as well as lengthy interviews also risk increasing errors of measurement, such as construct irrelevance. Finally, reported levels of significance are undoubtedly important, but useless, unless the magnitude of the difference is also reported.

Scores on vocabulary recognition tests also need to be thoroughly examined, for example: guessing possibilities undermine the reliability of true observed knowledge; an incorrect answer does not necessarily imply that there is a lack of knowledge of the target word but maybe of the matching alternative word (Warnby et al., 2022); a high reliability coefficient, such as Cronbach's alpha, does not guarantee high validity. However, the statistics help to detect possible validity problems and such issues were explored in Study I to a certain extent.

In this thesis, the variable under investigation is written receptive English academic vocabulary. Depending on the list of reference, these words amount to between 570 word families (AWL) and 3,015 lemmas (AVL). That is a small number of words in comparison to all the words in English in academic texts. Estimates of how many words one has to know in order to be able to cope with reading texts differ (see section 3.3.), but research in English has suggested that a learner who knows the 8,000–9,000 of the most frequent words will be able to read and understand texts of different types (Nation, 2006; Schmitt et al., 2011). A test of academic vocabulary will not be able to ascertain a connection between a cut-off score and academic reading success. However, a test of academic vocabulary can indicate if a learner lacks knowledge of the indispensable lexis of academic vocabulary. Not knowing the academic words will form a language barrier for reading and understanding texts at tertiary level (Edgarsson, 2018; Busby, 2020a).

6.7. Ethical considerations

Four main ethical requirements for properly conducted research are related to information, consent, confidentiality and usage (Vetenskapsrådet, 2017). All these requirements were met. Students were given relevant information about the research both orally and in writing, and they were given the opportunity to ask questions both before and after the data collection. Students themselves decided whether to participate and could retract their consent after the data collection. Participants were ensured that their anonymity was guaranteed, and that the data would be stored securely. Furthermore, the participants were informed that the data were only to be used for research purposes.

This study used a background questionnaire to collect information about students' L1 and country of birth, as well as their parents' L1 and country of birth. An ethical vetting procedure was made prior to data collection since such questions may be used to categorise and discuss students in terms of ethnic backgrounds[22]. The research was approved by the Swedish Ethical Review Authority. Ethnic background is not in focus in the thesis. Rather, variables such as L1 and country of birth could be used to investigate, for instance,

[22] Registration number 2019-04470.

whether different L1s or countries of birth may explain students' written receptive English academic vocabulary in different ways.

During the pilot phases, the questionnaires were discussed and altered continuously. For example, several students wanted the questionnaire to distinguish between a biological aspect and a self-perceived aspect of gender. Clearly, this was ethically important for the students, although, in the thesis, there is no intention of using the self-perceived gender variable in the analyses. For this reason, respecting the participants' feelings and expressed opinions I decided to include both questions in the final questionnaire (Appendix C).

During the data collection, in some classes, another ethical issue regarding background information about students' parents was discussed. Several students reacted to the questions about their parents. "Why do you need information about my parents?", was a question often heard among these students. *Parent information* is not one of the seven grounds of discrimination according to the Discrimination Act (SFS 2017:1128). However, for these students, some strong emotions made them refuse to answer those questions, whereas responding to biological gender or L1 met with no objections. The questions about parents are the ones with the most missing cells in the data and, therefore, restriction and cautiousness have been applied when analyzing such background variables.

7. Summary of studies I – III

The first study investigated two measurement instruments and linked scores from the VLT-Ac and the AVT to identify a valid and reliable measure of written receptive English academic word knowledge for the target population and to offer a model for score comparison between the two tests. The second study aimed at mapping the knowledge and development of students' written receptive English academic word knowledge during upper secondary school by comparing AVT scores in a sample of students when they entered upper secondary mandatory English instruction (*entry-sample*) with another sample of students after completion of the final mandatory English course (*exit-sample*). Moreover, Study II correlated academic vocabulary knowledge with achievement (final English grades) and estimated the probabilities for different sub-samples (programmes, gender and English entrance grades) of reaching a suggested vocabulary mastery threshold. In the third study, a special focus was on students' extramural English (EE) involvement and how this involvement correlated with academic vocabulary knowledge. Study III also explored whether the variation in academic vocabulary knowledge could be explained by such influences as EE factors, having one or more L1s, gender, age, parental educational level and number of years of EFL instruction. In both Study II and Study III, an underlying interest was to understand what sources may be significant for developing written receptive English academic word knowledge. This chapter will summarise the findings of the three studies.

7.1. Study I: Linking two tests of academic vocabulary

The first study set out to link scores from the VLT-Ac with scores from the AVT in a single-group linking design (Kolen & Brennan, 2014). The aim was to offer a comparison model, that is, to understand what a score on the VLT-Ac corresponds to in terms of a score on the AVT, and vice versa. Such a comparison model can be relevant to use when exploring trends in academic vocabulary knowledge over time when the choice of the administered test has varied between the AVT (Form 2) and the VLT-Ac (Version 1). Additionally, despite the risk of a ceiling effect in the VLT-Ac scores (e.g., Edgarsson, 2018) and despite more modern academic vocabulary instruments, scholars who want to compare their "results to the results of other studies" may still

choose "to use the original VLT" (Skjelde & Coxhead, 2020, p. 5/20). For this reason, a critical analysis of the construct and how it behaves in the VLT-Ac and the AVT with the target population was also reasonable. By using responses from 385 upper secondary students on the AVT and on the VLT-Ac an IRT-based concurrent calibration was conducted along with other statistical measures.

There are several ways of linking scores from different tests. Depending on the characteristics of the tests to be linked different procedures can be used. When two tests share the same framework and test specification the strongest form of linking, so-called equating, can be used. When two tests share the same overarching construct but have different test specifications, the scores from the two tests can be linked through a calibration procedure. The weakest linking procedure is known as statistical moderation, which is used when two tests which share neither the same framework nor the same specifications are administered to nonequivalent groups of examinees (Feuer et al., 1999; Kolen & Brennan, 2014). In Study I, it was argued both theoretically and empirically that the VLT-Ac and the AVT shared the same assessed construct, namely academic vocabulary, even if the test specifications were not identical (AWL vs. AVL, 30 vs. 57 target words) but the test formats were very similar. For this reason, the study adopted the second strongest form of linking, namely a concurrent calibration procedure using an IRT-based methodology in a single-group counterbalanced design.

In the data analysis, classical test theory was first used to explore the distribution of observed scores (Table 10).

Table 10

Descriptive statistics of AVT and VLT-Ac scores

		VLT-Ac	**AVT**
Mean		24.52	30.40
95% C.I. for Mean	Lower bound	23.99	29.25
	Upper bound	25.06	31.55
Std. Deviation		5.30	11.46
Standardised mean		.82	.53
Standardised *SD*		.18	.20
Mode		29	31
Minimum		3	2
Maximum		30	54
Quartiles	25^{th}	22	22
	50^{th}	26	31
	75^{th}	28.50	39

Note. Adapted from Warnby et al. (2023, p. 11)

The descriptive results (Table 10) revealed a clear ceiling effect for the VLT-Ac with this group of independent EFL users in expanding circles who had high exposure to English. The AVT scores, on the other hand, were normally distributed.

One of the key assumptions for adopting an IRT-based calibration is the assumption of unidimensionality, given that the two tests measure one major domain. An exploratory factor analysis revealed one major component within the two tests, thus strengthening the theoretical argument of one shared construct.

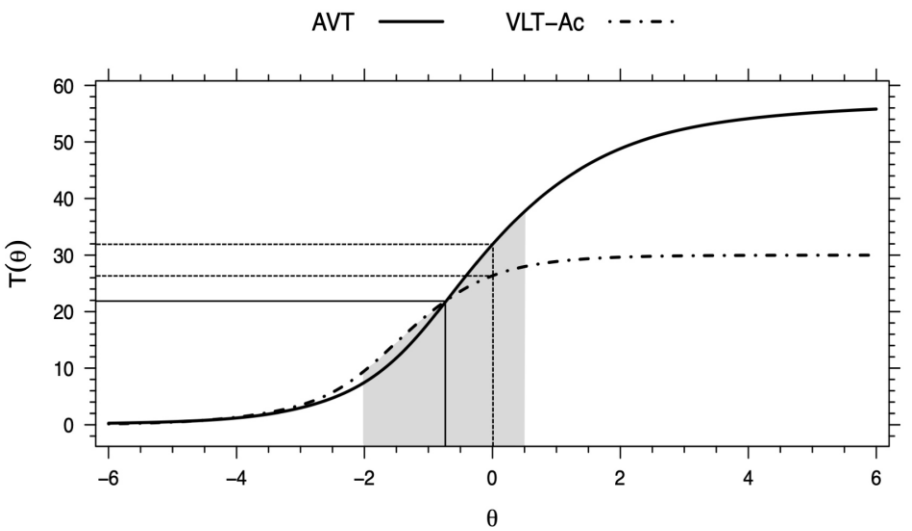
Several IRT models were explored initially (Rasch, 1PL, 2PL and 3PL) to check stability between estimated parameters and to inform on the decision of which IRT model to use for the main analysis and the calibration. The 2PL model was judged the best fitting model (Embretson & Reise, 2013, p. 184).

The estimation of parameters in the two tests revealed that the AVT items collected information on the full latent ability scale whereas the VLT-Ac lacked information at the higher end of the scale. The most reliable range of shared information between the two tests was from theta -2 to theta .5.

When all parameters from the 2PL model were checked for stability, a concurrent calibration was performed. During this procedure, all responses on the combined dataset (the two tests) were used to estimate all items in one single calibration run. Thereafter, the test characteristic curve (TCC) from each test was extracted from this combined set of estimates (Figure 16).

Figure 16

TCCs of the AVT and the VLT-Ac from the concurrent calibration



Note. Reproduced from Warnby et al. (2023, p. 14)

The TCCs in Figure 16 show, for example, that a score of approximately 26 on the VLT-Ac (y-axis) is located on theta = 0 (x-axis) and at this latent ability level a score of approximately 32 is expected on the AVT (y-axis). Hence, a test taker scoring 26 out of 30 on the VLT-Ac is expected to obtain an AVT-score of 32.

Since the shared information on the latent ability scale differs, especially at the higher end, and since there is a standard error to consider, a qualitative judgment of approximations was made to define the score comparison table, which constitutes the main result in Study I. This table included overlaps in scores at some levels. Thus, in the final score comparison table, a score of 26 on the VLT-Ac may vary between 31 and 33 points on the AVT (Table 11).

Table 11

Score comparison table

VLT-Ac Version 1	**AVT Form 2**
1–10	0–8
11–13	9–10
14	11
15	12
16	13
17	14–15
18	15–16
19	17–18
20	18–19
21	20–21
22	22–23
23	24–26
24	26–28
25	28–31
26	31–33
27	33–37
28	37–41
29	42–49
30	50–57

Note. Reproduced from Warnby et al. (2023, p. 15)

The score comparison table is the key finding and may serve as a practical tool for comparing themes such as academic vocabulary knowledge over time

where different instruments uses have been applied, or for predicting a student's score on one of the linked tests by using their observed score on the other test.

7.2. Study II: Measuring adolescents' academic vocabulary knowledge

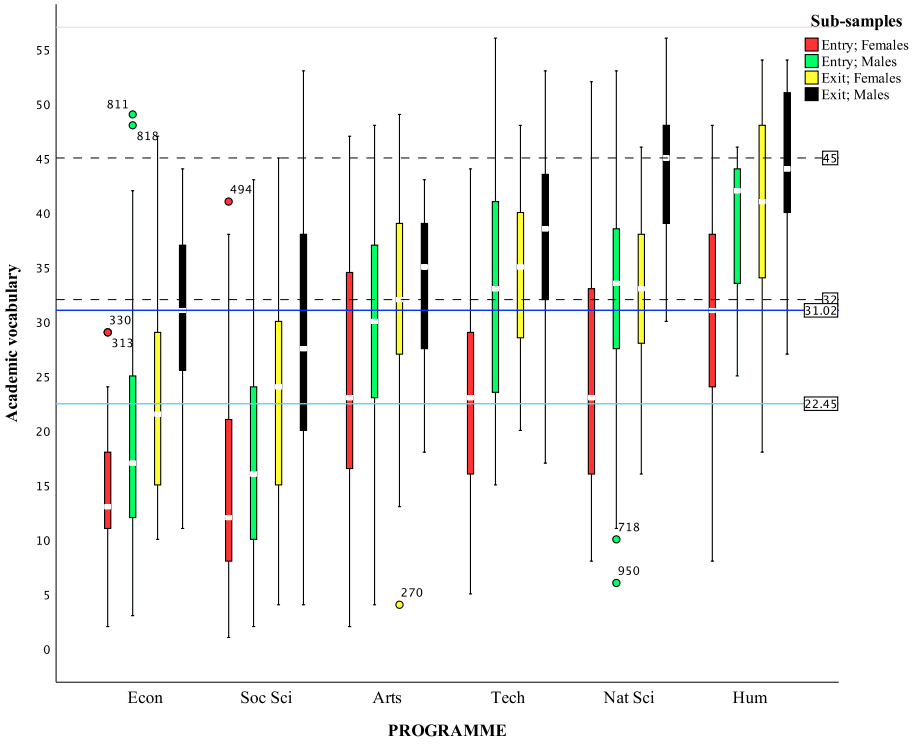
The second study set out to explore the written receptive English academic vocabulary knowledge exhibited (*i*) by students starting their first mandatory English course in upper secondary school (entry-sample, $n = 526$) and (*ii*) by students having completed their final mandatory English course (exit-sample, $n = 426$). Four RQs were asked:

> RQ1. What is the written receptive English academic vocabulary knowledge among upper secondary school students when they start and finish mandatory EFL instruction providing eligibility to higher education?
>
> RQ2. Are there any differences in academic vocabulary knowledge between first-year and final-year students with regard to gender and study disciplines?
>
> RQ3. Among final-year students, how are academic vocabulary scores and grades from the final English course correlated?
>
> RQ4. Do gender, study discipline and/or English entrance grades affect the probabilities for an upper secondary CEFR-B2 EFL school student to reach the suggested mastery thresholds of receptive English academic vocabulary?

For RQ1 and RQ2, descriptive statistics were used to display the academic vocabulary of the two samples as well as of their sub-samples (gender, study disciplines). Group mean differences between and within samples were tested for significance using T-tests. The boxplots in Figure 17 illustrate the variation in scores within and between sub-samples. As expected, the results revealed significant differences between the entry-sample and the exit-sample indicating that academic vocabulary knowledge develops during the two years. There were also large variations in AVT scores observed within samples and within sub-samples. Statistically significant mean score differences were observed between males and females. Large variations in AVT scores between study disciplines were also observed.

Figure 17

Boxplots of AVT scores separated by programmes, entry vs. exit sample and gender



Note. The light blue horizontal line indicates the observed mean score in the entry-sample, M = 22.45 (SD = 11.89); the dark blue horizontal line indicates the observed mean score in the exit-sample, M = 31.02 (SD = 11.66); The two dashed black horizontal lines indicate the lower (32 p.) and the higher (45 p.) mastery threshold. Adapted from Warnby (2023, p. 12).

For RQ3 and RQ4 only the exit-sample was used in the dataset. These students had all completed the final mandatory English course, English 6, required for entry to university. A pass grade from English 6 is said to be equivalent to CEFR B2. However, the grades in English 6 have a range of five pass categories, with E being the lowest and A being the highest. To answer RQ3, the English 6 register grades were correlated with AVT scores and a positive and moderate correlation of $\rho = .66$ ($p < .001$) provided additional support for the relationship between academic vocabulary and academic achievement (see Townsend et al., 2016).

RQ4 made use of previously suggested cutoff scores for the VLT-Ac, namely the widely adopted 86.6 percent mastery threshold, that is, 26 out of 30, and the higher 96.7 percent mastery threshold, that is, 29/30. The lower threshold has been used in similar studies, for example, in Iceland (Edgarsson, 2018) and in Norway (Skjelde & Coxhead, 2020). Reaching the lower threshold has been suggested to be a minimum of academic word knowledge for "basic comprehension" of reading academic texts (Edgarsson, 2018, p. 106). Scores beneath the higher threshold may be an indication to learners and teachers that a focus on academic vocabulary learning may be needed (Skjelde & Coxhead, 2020, p. 6/20). Since Study II used the AVT as the measurement instrument of academic vocabulary, the score comparison table presented in Study I was used. According to that score comparison table, the two mastery cutoff scores for the VLT-Ac (the lower 26/30 and the higher 29/30) corresponded to an AVT score between 31–33 and 43–47. A decision was made to use 32 as the lower corresponding mastery threshold and 45 as the higher in Study II. Two dummy variables were set. For the first dummy variable, all participants scoring at or above 32 points were coded as 1 and all beneath 32 points as 0. For the second dummy variable, all participants scoring at or above 45 points were coded as 1 and all beneath 45 points as 0. These dummy variables were then used as dependent variables in two separate logistic regression analyses. For each logistic regression model, the independent variables were three groups of study disciplines (Economics/Social sciences, Arts/Humanities and Technology/Natural Sciences), gender groups (females and males) and prior grades in English year 9 to account for differences when entering upper secondary education. The reference group in both models was a female student in Economics/Social science entering upper secondary education with the lowest pass grade in English. With statistical significance, the results showed an increased likelihood for a participant not belonging to either of the reference categories (female; Economics/Social science; lowest grade) to reach the mastery thresholds.

The overall conclusion was that despite the observed higher scores among final-year students in comparison to first-year students, many students exhibit low levels of academic vocabulary knowledge. The large variation in academic scores and the unequal odds of reaching the mastery threshold suggests that a substantial number of students lack an adequate level of academic word knowledge to read and understand academic texts; and, these results seem to correspond poorly with the curriculum intention of ensuring that *all* upper secondary students have sufficient knowledge to be well prepared for university studies, at least regarding reading academic English texts.

7.3. Study III: Academic vocabulary and extramural English

The third study explored sources for receptive academic vocabulary development with special attention to students' involvement in different extramural English activities, but also to the background factors age, gender, number of first languages, length of English instruction, and parental educational level. The dataset consisted of 817 upper secondary school students' responses to the AVT (all the students were 16 or 18 years old), an EE survey and background questionnaire. Previous research has shown positive correlations between extramural English involvement and general vocabulary knowledge, but this is the first study to specifically correlate pre-tertiary students' different EE activities with their written receptive English academic vocabulary knowledge.

The study posed three research questions:

> RQ1: To what extent are Swedish EFL upper secondary school students involved in different EE activities, and are there differences between age and gender groups?
>
> RQ2. How do students' different EE activities correlate with receptive academic vocabulary knowledge?
>
> RQ3. Which EE and demographic factors are significant predictors of students' academic vocabulary knowledge?

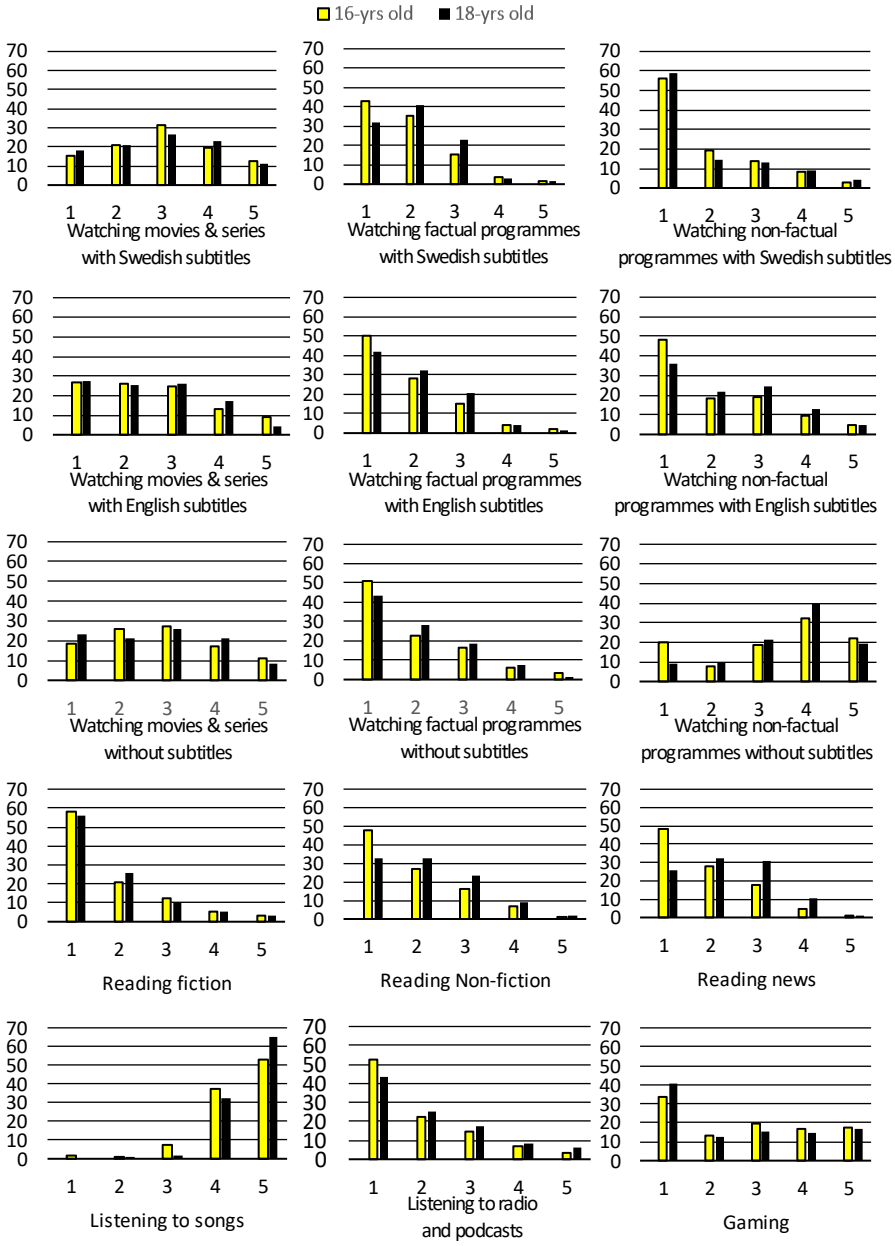
The AVT scores were used as a dependent variable in RQ2 and RQ3 with observed mean scores of M_{16yrs} = 21.7 (SD = 11.7) and M_{18yrs} = 29.6 (SD = 11.8).

For the first research question, the descriptive statistics confirmed what has already been shown before (e.g., Sundqvist, 2009), namely that Swedish adolescents are highly involved with extramural English activities. The bar charts in Figure 18 present the self-reported frequency of exposure and involvement with 15 different EE activities.

Furthermore, there were statistically significant differences in the amount of EE involvement between age groups; the older students were generally more involved in EE activities than the younger ones. However, as indicated by Cramer's V, the effect sizes were moderate or weak (Rea & Parker, 2014). Nonetheless, the pattern revealed a variation where, for example, the older students watched more English non-factual programs and read more English news and non-fiction than did their younger peers.

Figure 18

Bar charts of frequency of exposure to 15 different EE activities



Note. On the *x*-axis: 1 = /almost/never, 2 = once or twice/month, 3 = once or twice/week, 4 = /almost/ every day less than 2 hours, 5 = /almost/ every day more than 2 hours. On the *y*-axis (percentage): bars represent percentage of each sample group's reported extramural English involvement. Adapted from Warnby (2022, p. 132).

A similar pattern was found between gender groups, where, in general, male students were more involved in EE activities than the female students, except for reading fiction. The effect sizes were weak for all but one case, namely for the association between gender and gaming: males spend a lot more time playing English-mediated video games than do female students.

For the second research question, the findings in Study III revealed that several of the EE activities correlated positively with academic vocabulary, where reading fiction, non-fiction and news (printed or digital) showed the highest correlations. Only watching programs with Swedish subtitles had negative correlations. Although the correlations were weak, the results suggest that atypical academic activities such as gaming and watching non-subtitled programs offer possibilities for incidentally learning academic words.

Moreover, a linear model was performed to explore how academic vocabulary may be predicted from potential vocabulary acquisition sources. These sources were made up of the following factors: four composite EE activities (reading, viewing without subtitles, viewing with Swedish subtitles, viewing with English subtitles), age, gender, number of first languages, length of English instruction, and parental educational level. Gender and number of first languages did not show any statistical significance although males and students with more than one L1 had larger academic vocabulary knowledge than females and students with only one L1 respectively. The number of years of formal English instruction did not show any significance either. Age and parental educational level seemed to affect the academic vocabulary scores, but the effect sizes were tiny. The three composite EE factors (reading, watching without subtitles, and watching with Swedish subtitles) explained as much as 26 percent of the variance in academic vocabulary scores.

Thus, the results indicated that extramural English involvement seems to be a stronger source for academic vocabulary acquisition than any other tested factor, for example, length of English instruction. Although the lack of explanation from formal EFL instruction was a bit discouraging, it was not surprising since nothing in the curriculum directs instruction towards linguistic features of academic language such as academic vocabulary. Instead, the variation in academic vocabulary knowledge among these pre-tertiary students seemed to depend more on individuals' informal EE behaviors than formal university-preparatory schooling. Moreover, the higher involvement with gaming and non-subtitled programs (for example, YouTube) among males as compared to females is likely to affect their learning of receptive academic vocabulary.

One pedagogical implication put forward in the study was that EFL instruction could offer pre-tertiary students more activities that may support their incidental learning of academic vocabulary. A suggested possibility was, therefore, to include specific guidelines in the curriculum indicating the role of academic vocabulary and the possibilities to incidentally acquire this lexis, extramurally as well as intramurally, through extensive reading and viewing.

8. Discussion and reflections

> Jag skulle vilja känna mig mer förberedd för högskolan. Jag upplever att engelskan på gymnasiet inte fått mig att utvecklas så mycket som jag hade velat. Ex. kring kunskaper om struktur och ord i akademiska texter. När jag läser korta nyheter från The Guardian (på instagram) har jag svårt att förstå många ord, vilket gör det svårt att förstå kontexten. Upplever att dessa texter ofta är svårare än texterna vi haft i skolan.
>
> [I would like to feel more prepared for higher education. I feel that the English lessons at upper secondary haven't developed my English as I would have wanted. For example, knowledge about structure and words in academic texts. When I read short news items from *The Guardian* (on Instagram) I have difficulty understanding many words, which makes it difficult to understand the context. I believe those texts are often harder than the ones we've had in school.]
>
> – Julia, 18 years old, in her last upper secondary school year after final mandatory English course providing eligibility to university. From an open comment in the exit-sample questionnaires.

The student quoted above, from one of the samples, gives voice to opinions encountered many times during data collection. Furthermore, the statement provides a student perspective of the investigated construct in this thesis: written receptive English academic vocabulary knowledge. Key to the student's statement is the apparent lack of academic English instruction and the limited opportunities for engagement with academic English usage, potentially affecting the size of her academic vocabulary and, by extension, her ability to function in environments where knowledge of academic words will be critical. Her statement embodies the numerical results of this thesis, which will now be discussed and reflected upon.

The first section, 8.1, answers the thesis' overarching questions with reference to three aims investigated in the studies. Section 8.2. reflects upon these answers and findings in relation to the thesis' two central premises. Building on the reflections in 8.2., section 8.3. discusses pedagogical implications. Section 8.4. highlights some limitations of the thesis, and directions for future research are suggested.

8.1. Answering the thesis' research questions and addressing the aims

This thesis set out to answer two overarching questions:

RQ1: What is the written receptive English academic vocabulary knowledge among Swedish upper secondary school EFL learners?

RQ2: To what degree can specific educational and other individual factors explain the written receptive English academic vocabulary knowledge among Swedish upper secondary school EFL learners?

Three studies were conducted to answer these questions, and the aims of those studies were to a) identify a measure of the construct *written receptive English academic vocabulary knowledge* capable of yielding valid usage of scores pertaining to the construct and target population, b) map students' written receptive English academic vocabulary knowledge at the beginning of and the end of mandatory EFL instruction in upper secondary school, and c) explore potential factors of importance for written receptive English academic vocabulary knowledge.

The following numbered list highlights the contribution of the thesis in relation to the RQs and those three aims. Points 1–2 indirectly but inevitably address RQ1, numbers 3–5 speak directly to RQ1 and numbers 6–7 address RQ2 specifically:

[1] A factor analysis of scores on two tests (the VLT-Ac and the AVT) targeting words in the AWL and the AVL respectively revealed that the two tests measured one major component, which suggested that both lists and both tests capture a mutual overarching academic domain (Study I).

[2] The AVT seems to be a more appropriate measure of academic word knowledge than the VLT-Ac if administered to independent (CEFR-B2) users of English as a foreign language in a context with high exposure to extramural English (Study I).

[3] There is a large variation in academic vocabulary knowledge within the sample of final-year students (Study I and II), both within and between first-year and final-year samples (Study II) and between study programmes (Study II).

[4] There are students in all study programmes, of both genders, who exhibit a large academic vocabulary knowledge. Conversely, there are many students who exhibit a very poor level of academic vocabulary knowledge (Study II).

[5] In relation to previously suggested cut-off scores for mastery, approximately half of all students who passed the final mandatory English course (equivalent to CEFR-B2) do not reach the lowest suggested thresholds (regardless of whether the knowledge is measured with the VLT-Ac (Study I), or with the AVT (Study I and II).

[6] Extensive extramural English reading and viewing explain more of the variation in academic vocabulary scores than do years of formal English instruction, age, parental educational level, gender, and having one or more L1s (Study III).

[7] There are unequal probabilities for reaching mastery scores on the AVT for students in different study programmes and of different genders also when accounting for prior grades (Study II).

Taken together, the following general answers to the RQs can be made:

RQ1: Using scores from all three studies on the AVT[23], which performs as a suitable instrument for the target population, students entering upper secondary English instruction had an average score ranging between 21.7 to 22.45 (*SD* = 11.75–11.89) (Studies II & III); students who finished mandatory English instruction courses had average scores ranging from 29.63 to 31.02 (*SD* = 11.46–11.79) (Studies I – III). In more abstract terms, the answer to RQ1 is that there are widely varied patterns of written receptive English academic vocabulary knowledge among students who commence and students who leave mandatory upper secondary English instruction.

RQ2: Extensive extramural English involvement, gender and study discipline stand out as strong explanatory factors; parental educational level and age also explain the variation in academic vocabulary knowledge (Studies II & III). The strongest correlation found in Study III was between academic vocabulary knowledge and viewing programmes without subtitles on, for example, YouTube; however, the correlation was rather weak, $\rho = .43$ ($p < .001$). As a conclusion, the answer is that the potentially significant factors for developing a written receptive English academic lexicon to a large extent depend on /haphazard/behaviours/events on an individual/local level rather than on systematic events on an educational/global level.

However, it must be acknowledged that scores on a test are nothing more than scores on a test. What is interesting is when we want to extrapolate the scores to mean something substantial. The following section(s) draw on the answers and findings to discuss and reflect on their meaning from a broader perspective.

[23] A small reminder: the maximum possible score on the AVT is 57 points (see section 6.3.1.); building on the findings from Study I, the lower mastery threshold was set to 32 points and the higher mastery threshold to 45 point in Study II.

8.2. Reflecting on findings in relation to the thesis' two premises

As mentioned in 1.2., this thesis is based on two premises, namely (*i*) that academic vocabulary is an important facet of academic reading ability at university and, consequently, (*ii*) that academic vocabulary knowledge relates to the upper secondary curriculum objective of preparing all students equally for university studies. Section 8.2.1. reflects on the results in relation to the first premise whereas section 8.2.2. will review the second premise with reference to the results.

8.2.1. Average academic word knowledge level too low for academic reading comprehension

This thesis argues, as evidenced repeatedly in previous research, that reading comprehension depends, to a great extent, on vocabulary knowledge. Although vocabulary is the "single best predictor of text comprehension" (Alderson, 2000, p. 35), reading is a multi-faceted construct. Academic reading comprises, for example, vocabulary knowledge (general, academic and disciplinary), content knowledge, text type knowledge and metacognitive reading strategies. However, since vocabulary is the largest explanatory factor of reading comprehension, vocabulary researchers have tried to work out how many words in a text a reader must know to understand the text. At present, there is agreement that a lexical coverage of 98 percent offers the best likelihood that the reader will understand the text (e.g., Nation, 2001). The remaining unknown two percent of the words can be understood by drawing on contextual clues or the text comprehension will not be substantially affected by these two percent. Generally speaking, the words in academic texts consist of a large proportion of high-frequency words, 10–14 percent academic words (Therova, 2020), and, additionally, discipline-specific (usually low-frequency) words. Clearly, this thesis cannot establish whether students have a 98 percent lexical text coverage or not. Instead of speaking of upper secondary school students' future success in reading university course materials in English, the three studies offer a foundation for discussing whether there is a risk that students with very limited academic word knowledge are likely to be hindered in their future academic reading comprehension.

This thesis has shown that approximately half of the students in the exit-sample lack knowledge of a large number of academic words (i.e., a substantial proportion of the 10–14% of academic words in academic texts was unknown to them). This is a strong indicator that the academic reading comprehension of these students could be hindered (see section 4.3.). The lexical threshold will simply be too high. Consequently, the working memory is likely

to be occupied with struggling with word meaning (possibly including searching in a dictionary) to the detriment of building a model of text comprehension at a higher level. As regards the observed outcomes in this thesis, the validity argument for this statement builds on (*i*) the use of suggested mastery thresholds, (*ii*) the aspect of word knowledge tested in the instruments and (*iii*) the possible score inflation from guessing possibilities.

First, since the suggested mastery thresholds on the VLT have been contested, this thesis explored students' level of academic word knowledge according to the lower threshold as well as the higher threshold. By doing so, the scores can be discussed in more reliable ways. As noted in Chapter 1 and with reference to the curriculum and the English syllabi (see Chapter 2), this thesis is grounded in the premise that an appropriate level of preparedness for university studies (regardless of disciplinary orientation) is a certain level of academic vocabulary knowledge for general academic written receptive purposes. The mastery thresholds may thus be used as arguments for predicitive validity. The higher mastery threshold was used in Skjelde and Coxhead (2020) where only 8 percent of their Norwegian upper secondary school students reached this higher mastery threshold, which corresponds to the results found in Study II (11%). The lower mastery threshold level has been used in Germany (Coxhead & Boutorwick, 2018), Norway (Skjelde & Coxhead, 2020) and Iceland (Edgarsson, 2018) to investigate upper secondary school students' academic vocabulary knowledge as observed in the VLT-Ac. The positive correlation between VLT-Ac scores and IELTS academic reading scores (Edgarsson, 2018) indicates the valid use of academic vocabulary scores as an indicator of academic reading. On average, students below the lower mastery threshold had 55 percent or lower correct responses on the academic reading test which can be considered a weak IELTS result. The Swedish students in this thesis exhibit, on average, a similar level of academic vocabulary knowledge as in Edgarsson (2018). The correspondence between the lower academic vocabulary threshold and the IELTS academic reading score, as evidenced in Edgarsson (2018), supports the indication in this thesis: that the average academic vocabulary knowledge among this thesis' sample of Swedish upper secondary school students may be too small to be considered sufficient for taking on academic reading tasks at university.

Second, as explained in Sections 3.2. and 5.1., the receptive meaning-recognition knowledge tested in the VLT-Ac and the AVT is a basic aspect of word knowledge. Moreover, these tests build on the principle of partial lexical knowledge, which means that more nuanced meanings of word knowledge are not taken into consideration. In fact, the idea behind this principle is that test takers who barely know the meaning of a target word should be able to match it correctly with its definition (see Schmitt et al, 2001, p. 59). Considering this very basic aspect of word knowledge, the generally low average level of academic vocabulary knowledge in this thesis exit-sample (Study I and II) be-

comes even more critical for coping with academic reading at university. Proponents of the meaning-recall format as more representative of the reading construct (see section 3.2.2.) than meaning-recognition may have opted for a meaning-recall measure. Given the meaning-recognition results in this study, a design using measures of meaning-recall knowledge would very likely have generated results that are even more critical (and possibly produce floor-effects in the data).

Third, another critical issue that may pertain to the claim that the average academic vocabulary scores observed in this thesis (Study I and II) are low is the guessing possibilities within the VLT-Ac and the AVT. Previous research has suggested that the observed scores on the VLT are inflated due to guessing and, possibly, to exclusion strategies (e.g., Gyllstad et al., 2015). If the observed scores in this study are inflated, for example, by 17 percent as suggested by Stewart and White (2017, see section 3.2.2.), it would mean that the true knowledge is even lower than the observed, that is: an observed mean AVT score of 31.02 (Study II) could, in fact, be calculated as 31.02/1.17 = 26.51. Thus, such reasoning would critically increase the number of students who score far below the suggested mastery thresholds.

8.2.2. Large academic word knowledge variation in relation to curriculum objectives

The Swedish national curriculum postulates that the education must be based on principles of equality (SFS 2010:800). According to the results in this thesis, the large variation in academic vocabulary knowledge is, however, an indication that upper secondary school students who leave formal university-preparatory English instruction have very unequal academic vocabulary knowledge. This thesis has shown that male students have a higher probability of reaching sufficient academic vocabulary knowledge than their female peers (Study II). Moreover, this thesis has shown that male students are more involved in EE activities than their female peers – especially gaming and viewing without English subtitles – and that these male-dominant EE activities correlate positively with academic vocabulary knowledge (Study III). Therefore, the male students' higher probability of attaining the mastery thresholds may depend on their higher frequency of EE involvement. To the best of my knowledge, the statistically significant and positive correlations between EE involvement and academic vocabulary knowledge have not been evidenced earlier. It is encouraging that students' leisure activities can foster academic vocabulary growth, especially since this vocabulary construct is given little if any attention in school. From an educational assessment perspective, it is interesting that EE activities explain more of the variation in academic vocabulary knowledge than parents' educational level, gender and age, and, perhaps somewhat discouraging to find that the number of years of EFL instruction

revealed no statistically significant explanation (Study III). However, it is not surprising that knowledge that is excluded from the instructional content in the curriculum cannot be explained by education, but rather depends on individual differences. These results raise questions about the intention expressed in the Education act that schools shall counterbalance differences in all students' preconditions to absorb the education (see Chapter 2), for instance, absorbing basic knowledge to be prepared for university.

This thesis assumes that one aspect of the curriculum objective requiring that all students have "sufficient knowledge to be well prepared for studies in higher education" (NAE, 2013, p. 8) is to have a basic reading ability in academic English, which in turn, demands a certain level of written receptive academic vocabulary knowledge. The intended curriculum, however, does not state what this "sufficient knowledge" might mean or how it should be operationalised within the English subject. In fact, the English syllabi do not mention any kind of *academic English* nor *academic vocabulary* (not in the former curriculum, Gy11, nor in the current curriculum, Lgy22). The intended curriculum objective of "sufficient knowledge to be well prepared for studies in higher education" (NAE, 2013, p. 8) can thus be seen as very open to interpretations at a local level.

Although this specific thesis lacks data on how teachers interpret and enact the curriculum there are some other studies in the Swedish context describing vocabulary school practices in light of the curriculum. For example, Bergström et al. (2022b) interviewed secondary school teachers and found that, despite the teachers' general awareness of the importance of vocabulary for language performance, they did not view vocabulary as justified learning content to be taught and assessed according to the curriculum. Instead, the teachers relied, to a large extent, on incidental vocabulary learning. According to the teachers, words were expected to be, "picked up along the way" (p. 1) by the students through meaning-oriented activities. Moreover, the teachers also reported a lack of pedagogical approaches for initiating work that may foster the development of different aspects of word knowledge. When specifically asked to provide an account of instances where they did include vocabulary as part of their instruction, the teachers said 1) that they relied on their intuition and own experience about which words might be useful and 2) that the only specific vocabulary items they could think of in their instruction were "cohesive markers" (e.g., *also, therefore, although, especially*) (Bergström et al., 2022b, p. 13). This finding is particularly interesting given that cohesive markers were the only explicitly mentioned vocabulary content in the curriculum at the time of the interviews, that is: the 2011 curriculum (Gy11). Almost all teachers referred to the curriculum during the interviews which led Bergström et al. (2022b) to conclude that the teachers' "reported beliefs … appear to point to a possible curricular influence. This, in turn, raises the question to what extent the teachers are influenced by the absence of vocabulary and vo-

cabulary demands in the syllabus for English" (p. 13). The intended curriculum and the enacted curriculum seem interrelated as regards (the absence of) vocabulary instruction.

In another study using data from 1400 students, Eriksson (2021) found that about half the students said that the transition from upper secondary school to university studies was difficult with regard to reading in English, that they had negative feelings about English reading and that they were basically never instructed in reading strategies during upper secondary schooling. With regard to vocabulary and academic reading, a concern was that "when confronted with a word in English they [did] not understand, as many as 10 percent of university students [said] they give up and stop reading" (Eriksson, 2021, n.p.). Eriksson argued that the negative emotions (fear, anxiety, panic and stress) were a consequence of the upper secondary education's lack of preparing the students properly for English academic reading. What is expressed by the teachers in Bergström et al. (2022b) and by the students in Eriksson (2021) add to the information provided in this thesis' three studies and together they make up a broader empirical foundation for discussing the curriculum.

The English curriculum (i.e., the English syllabi) in the Swedish education system builds on a communicative approach, where language ability is strongly believed to develop as a result of meaning-oriented learning activities. Communicative language teaching (CLT) is characterised by its focus on successful rather than accurate language, on authentic rather than contrived materials, and on analysis of learners' needs rather than analysis of language. However, the theoretical broadness of the communicative approach has led to innumerable interpretations of CLT (see section 3.6.). Given the "arbitrary and ambiguous" (Siegel, 2019, p. 275) choices of language in the English communicative syllabi in Sweden and the "rather fuzzy nature of CLT … research is necessary if we are going to gain a better understanding of what learners exposed to a communicative approach to language teaching are actually learning" (Graves & Garton, 2017, p. 475). This thesis can be related to this call for research by questioning the curriculum objective that all students gain "sufficient knowledge to be well prepared for higher education" regarding written receptive academic vocabulary knowledge. By relating this thesis' findings to previous findings about Swedish secondary schools, for example, Bergström et al. (2022b) and Eriksson (2021, 2022), as well as to previous theoretical discussions about the *fuzziness* of CLT-based curricula, the validity of the arguments that can be made in this thesis are strengthened.

Hence, from this section, it can be argued that the majority of upper secondary school students in university-preparatory programmes do not leave English instruction with sufficient academic word knowledge to be well prepared for reading in English in higher education. Furthermore, it can be argued that there are unequal distributions of this knowledge within and across programmes even though all students follow the same English mandatory

courses. Such an arguably discouraging picture merits reflection upon possible pedagogical implications.

8.3. Pedagogical implications

This thesis has investigated academic vocabulary knowledge among upper secondary school students preparing for university and found that many students fall below the suggested thresholds of mastery of written receptive academic vocabulary knowledge. Furthermore, written receptive academic vocabulary knowledge has been used as an indicator of one facet of students' English academic reading ability. However, since academic reading contains other important skills, academic vocabulary is not to be regarded as the sole instructional content for teaching English for academic purposes. Instead, the pedagogical implications this thesis speak to concern the guidelines in the national curriculum.

Overall, in the steering documents concerning English instruction in upper secondary programmes offering preparation for higher education, there is an urge to specify that the English reading materials at university are demanding something else in terms of English proficiency; proficiency in colloquial English communication, a forte of many young Swedes, may not be sufficient. Arguably, there is a lack of constructive alignment between the overarching curriculum objective of preparing students for university studies and the instruction guidelines within the English syllabi (see sections 2.1.3. and 2.2.1.). This may very well be the case for many subjects, but this thesis has specifically looked into the English language demands. We cannot expect that the vague intention stated in the overarching curriculum objective will be equally enacted in English classrooms if the English syllabi do not include any guiding principles as to how this objective may be realised within the English subject.

We know that academic literacy is "largely not explicitly taught, but rather acquired indirectly, mostly through the practices in various school subjects, but also from practices out of school" (Prince et al. 2021, p. 176). The only specific reference to academic English reading skills in the steering documents is within the Diploma goals for the Technology programme where English is mentioned as important for taking part in contemporary technology research and innovations. For this reason, it would be valuable to investigate whether there are any differences in English literacy practices in various school subjects between the Technology programme and other programmes. Perhaps the higher probability of attaining academic vocabulary mastery for Technology students in relation to Economics/Social science students may depend on such under-researched issues. However, this thesis has shown that the patterns of academic vocabulary knowledge variation in Technology are similar to those in other programmes, especially with respect to the Arts, the Humanities and the Natural science programmes.

Thus, the first pedagogical implication that arises, I would argue, is a curricular change: a revision of the English syllabi and their adjacent commentary materials for the university-preparatory programmes is needed; a revision stating the importances of developing academic English proficiency, especially academic reading literacy, to be well prepared for taking on tasks already at the onset of tertiary level studies. Such academic English reading literacy consists of many facets, where academic vocabulary is one essential facet. This means that the English courses and syllabi for the university-preparatory programmes might need to contain somewhat different guidelines from those for the vocational programmes. Such differences already exist when it comes to, for example, the courses in mathematics and history in the first upper secondary year. Students in the Natural science programme take the course Mathematics 1c, students in the Social science programme take Mathematics 1b, and students in the vocational Healthcare programme take Mathematics 1a. These mathematics-1-courses all give 100 credit points, but the syllabi have differently expressed core content of instruction in relation to the study profiles. For History, the 1a courses are for students in vocational programmes whereas students in university-preparatory programmes are required to take History 1b. I argue that a similar logic may be suggested for the English syllabi, English 5 and English 6. This means that students in vocational programmes would take English *5a* and English *6a* courses whereas a university-preparatory programme student would have to take English *5b* and *6b* with a certain focus on teaching and learning English for academic purposes[24]. Furthermore, since the English syllabi are based on a communicative approach to language teaching and the syllabi contain no "explicit guidance about timings and priorities" (Siegel, 2019, p. 267), it is suggested that the syllabi could include more of a lexical approach to direct teachers to central aspects of academic English. A communicative language curriculum content that does not even identify the vocabulary registers required for communication can never help learners to develop into communicative language users (Milton, 2022).

Communicative competence demands a communicative lexicon (see section 3.5.); and the curriculum and expected learning outcomes are the first places to begin when outlining such a lexicon. Although rare, specific skills and knowledge areas expressed in the grading criteria do exist elsewhere in the Swedish education system, for instance, a student must be able to swim at least 200 meters to have a pass grade in Physical Education in lower secondary

[24] At the time of writing, the Swedish NAE (Skolverket, 2022e) is reforming the subject curricula for the upper secondary education. The reform focuses on the abolition of course grades in favour of subject grades, meaning that instead of getting two course grades – one for English 5 and one for English 6 – a student will leave secondary school education with one single subject grade in English. The reform suggests that the former 'courses' should be labelled 'nivå' ('level'). For mathematics the reform maintains the differences between a-, b- and c-levels, (Nivå 1a, Nivå 1b, Nivå 1c). The thesis' suggestion is the same regardless of label.

school, while a student must be able to give an account of the declaration of human rights to have a pass grade in the course Social Science 1 in upper secondary school. In line with such demands, one could argue that a student leaving lower secondary school should know at least the 2K band and a student leaving a university-preparatory programme should have knowledge of academic vocabulary. However, under-researched pedagogical ideas need to be well tested, assessed and evaluated before any curricular changes can take place.

A second implication, that follows as a consequence of the first, concerns what a well-balanced language course might look like, if based on a communicative approach and including an increased focus on lexis. Since academic reading comprehension demands a large lexicon, it is unreasonable to propose that English class time (≈ 460 hours in compulsory school and ≈ 180 hours in upper secondary) be used for the explicit instruction of all these words. A principled teaching approach may, thus, be required which takes into account incidental as well as intentional vocabulary learning opportunities. Since vocabulary learning builds on both quantity and quality, including frequent encounters with words and meaningful and adequate engagement with words, a possible solution may be to allot classroom time for extensive exposure to academic words in meaningful activities along with formal instruction of academic lexis (see Green, 2022). Study III clearly showed how extensive viewing and reading may be beneficial sources for academic vocabulary acquisition. Such vocabulary acquisition activities have previously been advocated by, for example, Krashen (1989) and Webb (2020) (see section 3.4.). However, until now, previous studies have not made the connection between receptive English academic vocabulary and extensive English exposure. Within Nation's four language teaching strands (1996, 2007), it is suggested that classroom instruction is divided into four equally large chunks of time: 1) meaning-focused input, 2) meaning-focused output, 3) formal language teaching, and 4) fluency. The first and the third strand are the ones this thesis mostly speaks to.

Meaning-focused input is often understood to refer to authentic materials within CLT. However, I would argue, that authentic materials (a tricky concept) are not construed for L2 learning research nor for pedagogical (i.e., school) purposes. Reading materials within the school context, for example, textbooks, should strive for a systematic approach to vocabulary development (see e.g., Bergström et al., 2023; Nordlund & Norberg, 2020) which increases the vocabulary load incrementally with respect to frequency and academic vocabulary as a basis for building academic English proficiency. With such a systematic way of producing reading materials (e.g., graded readers) learners are more likely to meet the lexical threshold, strengthen their vocabulary with repeated encounters and perhaps handle unknown words by using metacognitive reading strategies.

The formal language teaching strand would include instruction that highlights typical features of academic lexis. This thesis has shown that many students lack awareness of the most basic aspect of academic single-word knowledge. This would of course be a relevant content of instruction for formal language teaching; however, different learners have different needs. Instruments of the kind used in this thesis are easy to administer to students for diagnostic purposes and to find out where learners may have challenges. Other tools for measuring academic lexis (e.g., academic collocations or academic formulaic sequences) may provide additional information to teachers and learners on what to focus on in the teaching and learning of English academic lexis.

8.4. Limitations of the thesis and future research

With regard to this thesis' construct and to the aim of mapping students' academic word knowledge, there are three main limitations pertaining to the vocabulary tests. First, both the VLT-Ac and the AVT are small tests in relation to their underlying lists, the AWL and the AVL respectively. Second, the mastery thresholds for the VLT-Ac and the AVT need more empirical evidence for their accuracy to be verified. Third, the academic word knowledge assessed in this thesis, that is the aspect of meaning-recognition of written receptive academic vocabulary knowledge, is limited insofar as it does not fully represent the academic lexis required for successful reading. Future studies could opt for administering both versions of each test to increase the number of items and increase the reliability of the inferences that can be made about test takers' written receptive English academic lexicon. A future study on all the items in, for example, the AVL using a method such as the one used by Gyllstad et al. (2020) could provide more information about how many words need to be tested to better estimate mastery of academic vocabulary. Since reading academic texts also includes other aspects of knowledge of academic lexis, future studies are encouraged to include additional instruments measuring, for example, academic multi-word units, in their test battery (see Read & Dang, 2022).

Since much of the discussion in this thesis centres around academic vocabulary and academic reading, a limitation of the thesis is its lack of tests of academic reading comprehension. Instead, the thesis utilises evidence about the relationship between academic vocabulary knowledge and academic reading comprehension from other contexts as, for example, the study from Iceland where Edgarsson (2018) analysed and compared VLT-Ac scores with IELTS academic reading scores among upper secondary school students. Future studies in Swedish contexts could follow up on this topic by investigating,

for instance, the relationship between academic vocabulary mastery thresholds, academic IELTS reading score thresholds and the course pass grades from English 6 (CEFR B2-level).

Another limitation of the thesis regards the reliability of self-reported data and the generalisability of the target population. As mentioned in section 6.7, there are considerable gaps in the available information about parents. Register data from Statistics Sweden, for example, data on subject grades, GPA and parental educational level, would decrease the number of missing values or reduce risks of unreliable self-reported data.

Although approximately 500 participants per sample corresponds well to the sample size used in the construction of national assessment of foreign languages (Erickson et al., 2022) a larger sample size would increase the external validity, that is whether the results can be generalised to the Swedish upper secondary school student population preparing for university. Related to this point, a limitation of the samples is that none of the participating schools offered all six university preparatory study programmes, making it hard to distinguish between programme effects and possible school effects. Future studies are encouraged to use a larger and well-balanced sample size as well as register data to enable more elaborated analyses about patterns of variation in academic vocabulary knowledge.

As this thesis attempts to offer knowledge about the extent to which students in upper secondary education are prepared for reading texts in English from the assigned reading lists in higher education, the cross-sectional design including only data about academic vocabulary knowledge within upper secondary education is inevitably limited. Future studies may adopt a longitudinal design and/or use a mixed-method approach to strengthen the links between secondary and tertiary levels as well as between academic vocabulary knowledge and other variables related to academic literacy.

Finally, the pedagogical implications discussed in relation to this thesis' results and centred around the curriculum may be criticised for being too wide. It is therefore suggested that future studies might explore the relationship between the intended and the enacted curriculum, for example, by interviewing upper secondary school teachers about how they interpret and try to implement the curriculum objective of preparing students for higher education regarding the academic reading demands at university and by following and observing those teachers during classroom time. Another study could focus on the English literacy practices outside the English subject/classroom to see whether the upper secondary school practice prepares students for the English reading literacy practice in universities.

Some of the above-mentioned research suggestions are already ongoing and, hopefully, will be presented in the near future thus enriching the findings of this thesis.

9. Svensk sammanfattning (Swedish summary)

I det här kapitlet sammanfattas avhandlingen. Avsnitt 9.1. ger en beskrivning av och överblick över avhandlingens bakgrund, syfte och design. Avsnitt 9.2. ger en sammanfattning av de tre studierna som ingår. Avsnitt 9.3. redogör för kärnan i avhandlingens resultatdiskussion och de kritiska pedagogiska implikationer som den ger upphov till.

9.1. Bakgrund, syfte och övergripande design

På svenska universitet och högskolor är idag i genomsnitt hälften av kurslitteraturen skriven på engelska (Malmström & Pecorari, 2022). Att läsa och förstå akademisk text på engelska kräver flera olika kompetenser, men ordkunskap är en nyckelkomponent för läsförståelse (t.ex. Gough & Tunmer, 1986; Laufer & Ravenhorst-Kalovski, 2010; Qian, 1999; Schmitt et al., 2017).

Engelska akademiska texter kan grovt sägas utgöras av tre olika typer av ordförråd: allmänt, akademiskt och ämnesspecifikt ordförråd (t.ex. Nation, 2001; Coxhead, 2016). Det allmänna ordförrådet kallas ’allmänt’ för att det är frekvent förekommande i de flesta texttyper, till exempel *that, make* och *slow*. Ämnesspecifikt ordförråd är huvudsakligen termer som uteslutande hör till ett snävt ämnesområde eller disciplin, till exempel *molecule*, *polymer* och *oxygene* (Valipouri & Nassaji, 2013). Akademiskt ordförråd kan definieras som de ord som är typiskt kännetecknande för akademiska texter utan att vara ämnesspecifika eller alltför allmänna, till exempel *enable, proportion* och *subsequently*.[25]

Med avseende på ordkunskap anses generellt en läsare behöva kunna 98% av orden i en text för att förstå texten (Nation, 2001) och akademiska ord utgör ungefär 10–14% av engelska akademiska texter (Coxhead, 2000; Gardner &

[25] Denna indelning av ordförrådet i akademiska texter är även vanlig när vi talar om svenska texter. I Sverige har svenskt akademiskt ordförråd definierats och/eller undersökts av flera forskare (t.ex., Enström, 2004; Lindberg, 2006; Lindberg & Johansson, 2019; Ribeck et al., 2014; Warnby et al., 2022). Exempel på ord från en svensk akademisk ordlista (ESAO; Ribeck et al., 2014) är *baserad, förutsättning, huruvida,* och *medföra*.

Davies, 2014; Therova, 2020).[26] Stycket nedan – Exempel (1) – illustrerar förekomsten av akademiskt ordförråd i engelskspråkig kurslitteratur på grundivå. Textstycket är hämtat från Coe et al. (2019, s. 33) som läses som obligatorisk kurslitteratur i den första kursen i lärarutbildningen på Göteborgs universitet (se kapitel 4.3. för mer information). För att visa hur obegriplig texten blir utan kunskap om akademiska ord har dessa ord ersatts med nonsensord i Exempel (1).

> (1) A key immength here is that hing-broat memory is not just a storage portacity, moragious to an encyclopaedia or canita triffable on the internet; nor is it hebated to engeen facts. Instead, the paisture and saminitions among horsiments of memorised lictadge are teegily what bonate it to be used in solving remlods or genorbing samsite molks: if it is not paistured and obivible in memory, it cannot be used. Samporily, if a student has a good store of well-paistured lictadge, and fluent, automated skills, absorbing new ideas and forbediures is much easier.[27]

Stycket innehåller 92 löpord varav 26% är akademiska. Utan kunskap om dessa akademiska ord når läsaren endast 74% ordtäckning vilket är långt ifrån 98%-tröskeln (Nation, 2001). Om en läsare kan två tredjedelar av dessa akademiska ord räcker det endast till 91% ordtäckning. För att nå 98% ordtäckning kan inte mer än ett till två ord vara okända.

Engelsk akademisk ordkunskap är alltså relevant för läsandet av akademiska texter. Detta gäller oavsett studieorientering. Av det skälet har den här avhandlingen undersökt engelsk akademisk (och inte ämnesspecifik) ordkunskap hos gymnasiestudenter i samtliga sex nationella högskoleförberedande program. Dessa studenter får en bred behörighet till universitet och högskola och kan därmed i stor utsträckning byta studieinriktning på universitetet i förhållande till det specifika program de läst på gymnasiet.

Enligt den svenska gymnasieskolans läroplan är det skolans ansvar att varje enskild student ”på ett nationellt högskoleförberedande program inom gymnasieskolan ges möjlighet att uppnå kraven för en högskoleförberedande examen som innebär att eleven har *tillräckliga kunskaper för att vara väl förberedd för högskolestudier*” (Skolverket, 2011, Kap 2.1 Kunskaper, *Mål*) [min

[26] 98%-gränsen är inte absolut för den lexikala texttäckningen, men inom vokabulärfältet är man samfälligt överens om att kunskap av mellan 95-98% av orden i en text krävs för lyckad läsförståelse (Hu & Nation, 2000; Laufer & Aviad-Levitzky, 2017; Nation, 2001, 2013).

[27] Nonsensorden (här markerade med *) är i original: **immength – insight; *hing-broat –long-term; *portacity –facility; *moragious – analogous; *canita – information; *triffable – searchable; *hebated – limited; *engeen – routine; *paisture –structure; *saminitions – connections; *horsiments – elements; *lictadge – knowledge; *teegily – precisely; *bonate – enable; *remlods – problems; *genorbing – performing; *samiste – complex; *molks – tasks; *paistured – structured; *obivible – accessible; *Samporily – Conversely; *paistured – structured; *lictadge – knowledge; *forbediures – procedures.*

kursivering]. Gymnasieskolans obligatoriska engelskundervisning innebär att eleverna ska läsa Engelska 5 och 6 under sina första två år. Inget i ämnesplanerna nämner *akademisk engelska* eller *akademiskt ordförråd* som ett undervisningsmål som kan länkas till det övergripande målet att bli väl förberedd för högskolestudier. Likväl tydliggjordes relativt nyligen i den nationella lärarfortbildningssatsningen *Läslyftet* den centrala roll som ordkunskap har för läsförmågan.[28] I en av modulerna fick deltagande lärare läsa och utbilda sig i texten *Ordförrådet – en framgångsfaktor*:

> Expansionen av ordförrådet under skoltiden sker i stor utsträckning i samband med undervisning i de olika skolämnena som alla tillför ordförrådet sitt speciella språkbruk med tillhörande facktermer. Ett viktigt tillskott sker också genom ett mer allmänt skolspråkligt ordförråd som spelar en allt viktigare roll för lärandet ju högre upp i årskurserna eleverna kommer. Det är ett ordförråd som kännetecknar språkbruket i all formell utbildning, liksom i mer formella sammanhang i samhället i stort. Det handlar ofta om abstrakta, ”akademiska” ord vars betydelse kan vara svåra att sluta sig till utifrån ordens form. (Skolverket, 2017, s. 3)

De olika modulerna skrevs av forskare och lärare inom akademin på uppdrag av Skolverket och skulle bygga på insikter från forskning inom fältet. Utifrån textutdraget förstår man att modulinnehållet är mer detaljerat än vad som gick att uttyda ur den dåvarande läroplanen (Lgy11). Med tanke på vad som framhävs i textutdraget är det förvånande att akademisk ordkunskap fortfarande inte finns med som ett centralt innehåll i ämnesplanerna i den nyligen implementerade läroplanen, Gy22. Ingen lärdom från denna forskningsinsikt tycks ha påverkat revideringen av ämnesplanerna från Lgy11 till Gy22.

Utifrån denna bakgrund utgår avhandlingen från två premisser, nämligen 1) att engelskt akademiskt ordförråd utgör en central komponent i akademiska texter i svensk universitetskontext, och 2) att läroplansmålet ”tillräckliga kunskaper för att vara väl förberedd för högskolestudier” innebär en grundläggande förmåga att läsa akademisk engelska. Tidigare forskning i Sverige som beaktat gymnasiestudenters engelska akademiska ordkunskap är begränsad, och ingen tidigare forskning har haft just skriftlig receptiv engelsk akademisk ordkunskap som huvudsakligt undersökningsobjekt och inte heller i relation till ovannämnda premisser. Av det skälet sökte den här avhandlingen besvara följande två frågor:

---

[28]”Läslyftet är en kompetensutveckling i språk-, läs- och skrivdidaktik baserad på aktuell forskning. Läslyftet i skolan vänder sig till lärare i alla ämnen i grundskola, grundsärskola, sameskola, specialskola, gymnasieskola, gymnasiesärskola, lärare och förskollärare i förskoleklass samt skolbibliotekarier” (Skolverket, 2023, avsnitt 11).

1) Hur omfattande är den receptiva engelska akademiska ordkunskapen hos gymnasiestudenter i svenska högskoleförberedande program?
2) Vilka utbildningsfaktorer och andra individuella faktorer kan förklara den receptiva engelska akademiska ordkunskapen hos gymnasiestudenter i svenska högskoleförberedande program?

För att kunna besvara dessa frågor genomfördes tre studier som avsåg att a) identifiera ett mätinstrument avseende skriftlig receptiv akademisk engelsk ordkunskap och vars resultat ska kunna användas på ett reliabelt och valitt sätt i relation till avhandlingens kontext och syfte, b) kartlägga den receptiva engelska akademiska ordkunskapen hos gymnasiestudenter i svenska högskoleförberedande program och c) undersöka vilka utbildningsfaktorer och andra individuella faktorer som erbjuder förklaringar till gymnasieelevers receptiva engelska akademiska ordkunskap. De faktorer som den senare analysen riktas mot är (*i*) avslutande betyg i engelska från grundskolan och från gymnasieskolans sista obligatoriska engelskkurs, (*ii*) studieprogram, (*iii*) kön, (*iv*) antal år av formell engelskundervisning, (*v*) antal förstaspråk, (*vi*) ålder, (*vii*) föräldrautbildningsgrad och (*viii*) användning av engelska utanför lärandemiljön, så kallad extramural engelska (Sundqvist & Sylvén, 2016).

Avhandlingen använder befintliga testinstrument för receptiv kunskap om akademiska ord såsom de definieras i befintliga korpusbaserade akademiska ordlistor. Testpoängen används som indikatorer på elevers förutspådda akademiska läsförståelse. Validitetsargumenten för att använda de valda akademiska vokabulärtesten är tre och redogörs för i avhandlingen men förenklas här:

1. Gemensamt för flertalet teorier kring läsning är att ordkunskap är en grundläggande och avgörande komponent i läsförståelseprocessen (t.ex. Gough & Tunmer, 1986; Grabe & Stoller, 2019).

2. Som sades inledningsvis gör estimeringar gällande att det krävs kunskap om 98% av löporden i en text för att förstå texten (t.ex. Nation, 2001). De återstående 2% antas förstås genom till exempel kontextuella ledtrådar, och/eller textens betydelse kan härledas utan kunskap om de orden. Oavsett texttyp är normalt 80% av löporden i en text högfrekventa. I akademiska texter är cirka 10–14 % av orden akademiska (Therova, 2020). Logiken är enkel: utan akademisk ordkunskap kommer läsaren således inte att nå 98 % lexikal texttäckning.

3. Den typ av ordkunskap som testas i avhandlingen ligger på en basal och grundläggande nivå, nämligen att koppla ett ords form till dess betydelse (s.k. *meaning-recognition*). Mer avancerade ordkunskapstyper kan till exempel innebära kunskap om ett ords flera betydelseskillnader (t.ex. polysemi) eller att själv kunna verbalisera ett ords betydelse (s.k. *meaning-recall*) (t.ex. Nation, 2001). Om en läsare saknar en sådan grundläggande kunskapsnivå som testas i avhandlingen (*meaning-recognition*) är det troligt att hen inte heller når mer nyanserade kunskapsnivåer (t.ex. betydelseskillnader, *meaning-recall*) som kan vara centrala för den djupläsning som akademiska texter kräver (Milton, 2009).

Utifrån dessa premisser och validitetsargument syftar avhandlingen till att utöka vår förståelse om den skriftliga receptiva engelska akademiska ordkunskapen hos studenter när de påbörjar och avslutar de obligatoriska engelskkurserna i gymnasieprogram som avser förbereda dem för högskolestudier. Vidare utforskar avhandlingen faktorer som kan ha statistiskt signfikant förklaringskraft till denna akademiska vokabulärkunskap.

Avhandlingen använder en tvärsnittsdesign med totalt 998 elever från samtliga sex nationella högskoleförberedande program: 552 elever testades när de precis började Engelska 5 (*entry-sample*) och 446 när de avslutat Engelska 6 (*exit-sample*). Sammanlagt deltog 44 klasser från såväl kommunala skolor som friskolor i fem olika kommuner i Mellansverige. Samtliga elever svarade på enkätfrågor och gjorde vokabulärtest.

Bakgrundsinformation som användes innehöll ovan nämnda faktorer, till exempel betyg i engelska, studieprogram och kön. Eleverna fick också på en femgradig skala ange hur mycket de uppskattade att de ägnade sig åt olika engelskmedierade fritidsaktiviteter, såsom till exempel att läsa engelsk skönlitteratur, facklitteratur och nyheter, se på svensktextade, engelsktextade eller otextade engelska program, lyssna på engelsk musik och engelska poddar och spela datorspel där engelska används. Frågorna tog inte hänsyn till om läsning avsåg tryckt text eller skedde online utan intresserade sig i stället för genre. Samtliga elever gjorde det engelska akademiska ordkunskapstestet *The Academic Vocabulary Tes*t (AVT) framtaget av Pecorari et al. (2019) som har en maxpoäng på 57 poäng. I *exit-sample* fick elever även göra den akademiska sektionen i det välkända testet *Vocabulary Levels Test* (VLT-Ac; Schmitt et al., 2001) som har en maxpoäng på 30 poäng. AVT och VLT-Ac har ett matchningsformat där eleverna ges tre definitioner/synonymer som sedan var och en ska matchas med tre akademiska målord från en lista av sex alternativ. Alla instrument genomfördes med papper och penna och all data samlades in med forskaren på plats i klassrummet. Forskningsdesignen har godkänts i etikprövning. All data matades först in i Excel och processades sedan i statistikprogrammen SPSS och/eller R. Såväl deskriptiv som inferentiell statistik användes.

9.2. De tre studierna

Här sammanfattas de tre studierna.

9.2.1. Studie I

I Studie I användes 385 deltagares svar på AVT och VLT-Ac. Det primära syftet – till vilket forskningsfrågan kopplade – var att länka poängen mellan testen. Bakgrunden till detta syfte ligger i att kritik har lyfts mot det äldre och väletablerade VLT samtidigt som det modernare AVT inte är utforskat i samma utsträckning. Trots befintlig kritik tenderar forskare att välja VLT-Ac för att kunna göra jämförelser med tidigare forskningsresultat (t.ex. Skjelde & Coxhead, 2020). Ett indirekt mål var följaktligen att kritiskt undersöka testen och det avsedda konstruktet – receptiv engelsk akademiskt ordkunskap. Som en naturlig följd härav undersöktes hur testen kan förstås när de administreras till engelska inlärare som förbereder sig för universitetsstudier i en kontext där de utsätts för mycket engelska på fritiden och där deras språkliga nivå klassas som den gemensamma europeiska referensramens (CEFR) B2-nivå.

Genom att använda *Item Response Theory* (IRT)-analyser (Embretson & Reise, 2013; Hambleton & Swaminathan, 2013) kunde en samtidig kalibrering göras och en poängtabell tas fram som länkar poängen mellan testen. På så sätt kan tidigare testresultat utifrån Version 1 av VLT-Ac jämföras med framtida resultat på Form 2 av AVT. Med hjälp av jämförelsetabellen kan även en individs resultat på det ena testet användas för att estimera samma individs resultat på det andra testet. Till exempel kan en testtagare som får 26/30 poäng på VLT-Ac estimeras få ungefär 32/57 poäng på AVT.

Ett annat resultat var att AVT framstod som ett mer lämpligt test på den avsedda målgruppen eftersom det testet lyckades göra estimeringar över hela skalan, medan VLT-Ac endast mätte i den lägre delen av den skalan. VLT-Ac uppvisade takeffekter med denna målgrupp, men kan likväl vara ett lämpligt test för en annan målgrupp.

9.2.2. Studie II

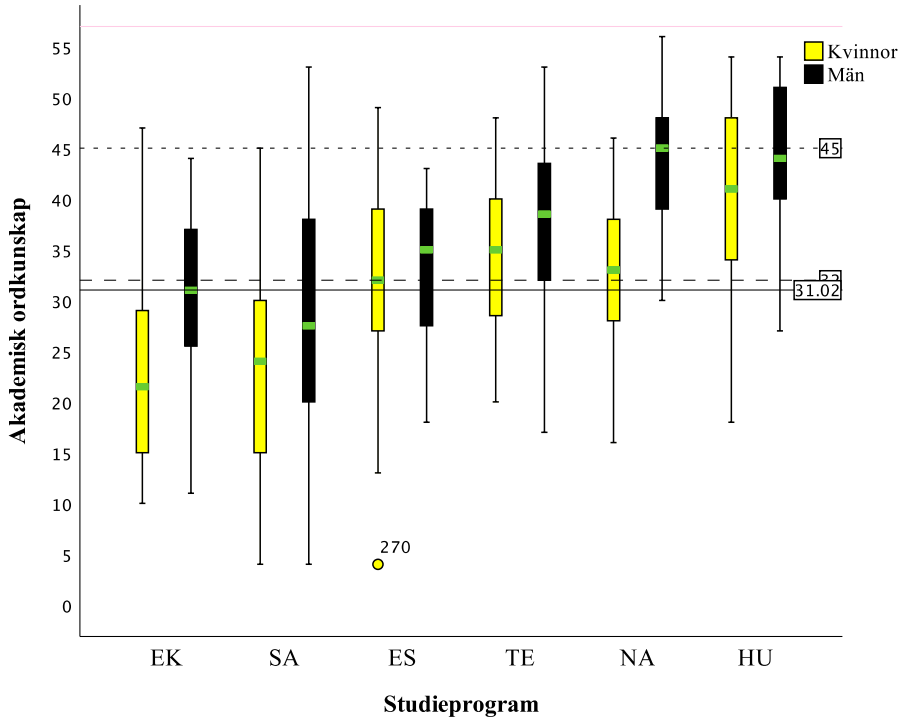
I Studie II undersöktes den akademiska ordkunskapen hos 952 elever varav 526 i *entry-sample* och 426 från *exit-sample*. I studiens första del undersöktes även kunskapsskillnader mellan *entry*- och *exit*-grupperna, mellan kvinnliga och manliga elever och mellan studieprogram. I studiens andra del låg fokus på eleverna i *exit-sample* som hade avslutat Engelska 6 med godkända betyg och därmed hade nått kravet för att ha behörighet till högskolestudier, åtminstone vad beträffar engelskundervisningen. Först undersöktes sambandet

mellan akademisk ordkunskap och betyg i Engelska 6. Därefter användes två av tidigare forskning föreslagna tröskelnivåer för att undersöka sannolikheten att nå dessa tröskelnivåer beroende på studieprogram, kön och tidigare engelskbetyg. Poängen som användes var 32 poäng för den lägre tröskelnivån och 45 poäng för den högre tröskelnivån (se även Studie I). Dessa tröskelnivåer användes som beroende variabler i två logistiska regressionsmodeller. De oberoende variablerna som användes var kön, studieprogram och betyg i engelska från årskurs 9. De kvinnliga eleverna utgjorde referensgrupp för könsvariabeln. Samhällsvetenskaps-/ekonomiprogram utgjorde referensgrupp för studieprogramsvariabeln. Betygen E/D utgjorde referensgrupp för betygsvariabeln. Anledningen till att använda engelskbetyg från årskurs 9 i modellen var för att kunna jämföra elevernas ingångsvärde innan de genomgått två år av gymnasial engelskundervisning.

Resultaten visade bland annat att det var statistiskt signifikanta skillnader med meningsfulla effektstorlekar mellan och inom tvärsnittsgrupperna, där, till exempel, *exit-sample* hade markant större ordkunskap än *entry-sample*. Likaså visade sig skillnaderna mellan studieprogram och kön vara statistiskt signifikanta. Lådagrammen i Figur 19 illustrerar den stora ordkunskapsvariationen mellan program och kön inom *exit-sample*.

Figur 19

Lådagram över den engelska akademiska ordkunskapsvariationen mellan program och kön hos exit-sample



Kommentar. EK = Ekonomi-, SA = Samhällsvetenskaps-, ES = Estetiska -, TE = Teknik-, NA = Naturvetenskaps-, HU = Humanistiska programmet. Heldragen horisontell linje indikerar medelvärdet för hela *exit-sample*, streckad linje anger det lägre tröskelvärdet 32 poäng på AVT och den prickade linjen det högre tröskelvärdet 45 poäng på AVT. Figuren är en svensk anpassning av den som presenterades i originalstudien (Warnby, 2023).

Ett linjärt samband visade sig mellan akademisk ordkunskap och betygen efter den sista obligatoriska engelskkursen, Engelska 6.

Slutligen visade de logistiska regressionsanalyserna att elever som inte tillhörde referensgruppen (kvinna; samhällsvetenskaps-/ekonomiprogram; betyg E/D från årskurs 9) hade betydligt större sannolikhet att nå föreslagna tröskelnivåer för tillräcklig akademisk ordkunskap (32 poäng för den lägre och 45 poäng för den högre tröskelnivån). Det betyder till exempel att det är markant troligare att en manlig elev i Samhälls- eller Ekonomiprogrammet når tröskelnivåerna på AVT än en kvinnlig elev från samma program givet att de båda

hade samma betyg när de påbörjade gymnasiet. Det är också betydligt mer sannolikt att en kvinnlig elev som gått i Teknik-, Naturvetenskaps-, Estetiska eller Humanistiska programmet når tröskelnivåerna än en kvinnlig elev som gått Samhälls- eller Ekonomiprogrammet givet att de båda lämnade grundskolan med samma engelskbetyg.

Sammantaget visade resultaten på stor variation i akademisk ordkunskap och ojämlika villkor för att sannolikt lära sig detta, vilket verkar bero på lokala och/eller individuella faktorer. Därutöver når ungefär hälften av eleverna som avslutat sin obligatoriska engelskutbildning inte tillräcklig nivå på AVT vilket kan vara en indikation på att cirka hälften kommer att få svårt att läsa och förstå akademiska texter på engelska i högre utbildning.

9.2.3. Studie III

Studie III fokuserade på 817 elevers (16-åringar från både *entry-sample* och 18-åringar från *exit-sample*) grad av exponering för fritidsengelska, så kallad *extramural engelska* (Sundqvist & Sylvén, 2016), i relation till akademisk ordkunskap. Tidigare forskning har visat att extramural engelska bidrar till en individs allmänna engelska språkfärdighet (t.ex. Sundqvist, 2009), men inga tidigare studier har utforskat eventuella samband mellan extramural engelska och akademisk ordkunskap.

Korrelationer visade att det fanns positiva samband mellan till exempel läsning av skönlitteratur och akademisk ordkunskap, mellan tittande på otextade engelska program (på till exempel YouTube) och akademisk ordkunskap, och mellan datorspelande (*gaming*) och akademisk ordkunskap. Sådana fynd ger en indikation på den möjlighet som till exempel *gaming* kan ha på inlärandet av receptiva akademiska språkfärdigheter. Exempel (2) och (3) belyser den typ av texter som spelare av *Call of Duty* (ett spel som nämndes av flera spelkonsumerande informanter) kan möta i olika undermenyer, först med samtliga akademiska ord ersatta med ett X, se exempel (1), och därefter så som texten står i original, se (2).

> (2) **Battle Pass X** – A tiered X X that offers X personalization and Loadout X. The Battle Pass X are free tiers that are X for all players, X of whether they own the Battle Pass or not. Purchase the Battle Pass to X X to unlock 100 Tiers of X.

Exempel (1) illustrerar hur obegriplig texten blir om man som läsare/spelare saknar förståelse av de elva utelämnade akademiska orden. Det blir samtidigt uppenbart i exempel (2) att exponering för den här typen av texter ger stora möjligheter att (mer eller mindre omedvetet) lära sig akademiska ord:

(3) **Battle Pass System** – A tiered content system that offers multiple personalization and Loadout items. The Battle Pass System are free tiers that are accessible for all players, regardless of whether they own the Battle Pass or not. Purchase the Battle Pass to receive access to unlock 100 Tiers of content.
(https://www.callofduty.com/warzone/strategyguide/welcome-to-the-warzone/In-GameTerms)

Vidare gjordes en regressionsanalys för att undersöka vilka faktorer som förklarade mest av skillnader i akademisk ordkunskap. Utöver faktorn extramural engelska introducerades ålder (16 versus 18 åringar), föräldrarutbildningsgrad (icke-akademisk versus akademisk bakgrund), antal förstaspråk (de med ett eller de med två eller flera), och antal år av engelskundervisning. Resultaten visade att extramural engelska förklarade 26% av skillnaderna i akademisk ordkunskap, vilket var långt över ålder (2%), föräldrautbildningsgrad (1%) och övriga undersökta faktorer.

9.3. Resultatdiskussion och implikationer

Avhandlingen avsåg huvudsakligen att undersöka den akademiska ordkunskapen hos gymnasieelever när de börjar och slutar obligatorisk engelskundervisning och identifiera vilka faktorer som erbjuder förklaringar till denna kunskap.

De sammantagna resultaten från de tre studierna avslöjar stora variationer i akademisk vokabulärkunskap såväl inom som mellan urvalsgrupperna. Viktiga faktorer som visar sig förklara den akademiska ordkunskapen hos gymnasieeleverna är extramural engelska, ålder, kön och studieprogram. Anmärkningsvärt är att i genomsnitt hälften av eleverna som lämnar den högskoleförberedande gymnasieskolans obligatoriska engelskutbildning inte når det föreslagna lägsta tröskelvärdet för vad som indikerar behärskning av receptiv akademisk ordkunskap.

De didaktiska implikationer som dessa resultat ger upphov till rör det uppenbara: eftersom det inte finns några riktlinjer i de engelska ämnesplanerna om akademisk vokabulärkunskap är resultaten förväntade. Står det inget i läroplanen om vikten av att kunna akademiska ord för att lyckas läsa akademiska texter kan vi inte heller förvänta oss någon form av tydlig och sammanhållen kunskapsbild avseende akademisk ordkunskap hos eleverna vid utbildningens slut. Den stora andelen elever som inte tar sig över den föreslagna lägsta tröskeln är med största sannolikhet inte tillräckligt förberedda för att ta sig an universitetens och högskolornas engelska läsuppgifter. Det vill säga, det är högst tveksamt om denna andel av elever har ”tillräckliga kunskaper för att vara väl förberedd för högskolestudier” (Skolverket, 2011, Kap 2.1 Kun-

skaper, *Mål*). Avhandlingen drar därför slutsatsen att en förändring i läroplanen krävs eftersom den konstruktiva länkningen (*constructive alignment*) mellan läroplansmålet och ämnesplanernas riktlinjer kan sägas saknas (Biggs, 2003). Slutsatsen ger upphov till två konkreta åtgärdsförslag vars strukturella logik är jämförbar med logiken som återfinns i andra ämnen på olika ställen i utbildningssystemet.

Det första åtgärdsförslaget bygger på logiken som återfinns i exempelvis matematikkurserna på gymnasiet. Elever i de yrkesförberedande programmen läser Matematik 1a (100 poäng) medan elever i till exempel det högskoleförberedande Samhällsvetenskapsprogrammet läser Matematik 1b (100 poäng) och elever i Naturvetenskapsprogrammet Matematik 1c (100 poäng). Utifrån en sådan logik kan man argumentera för följande åtgärdsförslag: att elever i yrkesförberedande program ska läsa engelskkurser som i likhet med matematikkurserna skulle kunna betecknas Engelska *5a* och Engelska *6a*[29] och att elever i högskoleförberedande program läser Engelska *5b* och Engelska *6b*. Ämnesplanerna i *5a* och *6a* skulle ha ett mer yrkesorienterat engelskinnehåll i förhållande till Engelska *5b* och *6b* som skulle ha ett mer högskoleorienterat engelskinnehåll. Förslagsvis skulle ämnesplanerna i sådana Engelska *5b*- och *6b*-kurser ha ett framskrivet innehåll kring akademisk engelsk läsförmåga som oumbärlig kunskap i högre utbildning och därmed länkas till det övergripande läroplansmålet.

Det andra åtgärdsförlaget bygger på logik som återfinns i utbildningssystemet gällande betygskriterier som beskriver konkreta ämneskunskaper som krävs för att få godkänt betyg. Ett första exempel är att i ämnet Idrott och hälsa krävs för det lägsta betyget E att eleven kan simma ”200 meter i en följd varav 50 meter i ryggläge” när hen slutar årskurs 9 i grundskolan (Skolverket, 2022h). Ett andra exempel är att i kursen Samhällskunskap 1b på gymnasiet krävs för betyget E att eleven besitter tillräckliga faktakunskaper så att hen kan ”redogöra för de mänskliga rättigheterna” (Skolverket, 2022d). Utifrån sådan logik skulle man kunna hävda att en elev som slutar årskurs 9 ska behärska de 2000 mest högfrekventa engelska orden och att en högskoleförberedande gymnasieelev som slutar Engelska 6 (eller företrädesvis *6b*) ska besitta kunskap om engelskt akademiskt ordförråd. Baserat på denna avhand-

[29] I skrivande stund pågår en remissrunda kring en reformering av ämnesplanerna i gymnasieskolan. Denna *ämnesbetygsreform Skolverket* (2022f) rör som titeln anger huvudsakligen ett övergivande av kursbetyg till förmån för ett ämnesbetyg. I stället för kurser, såsom till exempel Engelska 5, Engelska 6 och Engelska 7, innebär reformförslaget att ämnet delas in i nivåer, vilka etiketteras Nivå 1, Nivå 2 och Nivå 3 (vilka i stora drag är jämförbara med Engelska 5-6-7). För matematikämnet innebär reformen att även indelningen av nivåer i 1a, 1b, 1c, 2a, 2b, 2c osv. kvarstår. Resonemanget som förs i avhandlingen om en skillnad mellan Engelska 5a/6a för yrkesprogram och 5b/6b för högskoleförberedande program är givetvis lika giltig oavsett etikett (Nivå 1a/2a vs Nivå 1b/2b).

lings undersökning innebär alltså åtgärdsförslaget en sådan reformering av betygskriterierna för engelska i gymnasieskolan (företrädesvis i den ovan föreslagna kursen Engelska *6b*).

Ämnesplanen i engelska genomsyras av en kommunikativ språksyn som ibland kritiserats för att vara otydlig och som har lett till en mängd olika undervisningspraktiker (Butler, 2011; Graves & Garton, 2017; Stelma, 2009). Den här avhandlingen föreslår inte en ren och skär ordkunskapsundervisning utan förordar en balanserad undervisning, till exempel såsom föreslås i Nations *Four strands* (1996, 2007). Meningsfulla kommunikativa aktiviteter speglar ett autentiskt språkbruk bättre än ett instrumentellt fokuserande på till exempel enskilda språkliga egenheter. Men även om meningsfulla och autentiska kommunikativa situationer är målet med den kommunikativa språksynen betyder inte det automatiskt att sådana situationer/aktiviteter också alltid är den bästa undervisningspraktiken. Målet är inte alltid medlet. För att vara en kommunikativ språkbrukare behövs ett kommunikativt ordförråd passande den kommunikativa situationen (Milton, 2022), och det finns väletablerade beskrivningar av hur ett ordförråd behöver vara beskaffat beroende på kommunikativ situation. I akademiska lässituationer utgör akademiskt ordförråd en sådan oumbärlig beskrivning. En läroplan som i sitt centrala undervisningsinnehåll undviker sådana etablerade beskrivningar sviker såväl lärare som elever när det gäller att skapa en god förutsättning för att klara av att läsa engelsk kurslitteratur på universitetet och för att nå det högskoleförberedande läroplansmålet.

References

Abbasian, R., Hadian, B., & Vaez-Dalili, M. (2020). Examination of the role of family socio-economic status and parental education in predicting English as a foreign language learners' receptive skills performance. *Cogent Education*, 7: 1710989, 1–13. https://doi.org/10.1080/2331186X.2019.1710989

Ackermann, K., & Chen, Y.-H. (2013). Developing the Academic Collocation List (ACL)–A corpus-driven and expert-judged approach. *Journal of English for Academic Purposes*, *12*(4), 235–247. https://doi.org/10.1016/j.jeap.2013.08.002

Airey, J. (2009). *Science, language, and literacy: Case studies of learning in Swedish university physics*. Acta Universitatis Upsaliensöis.

Alderson, J. C. (2000). *Assessing reading*. Cambridge University Press.

Alderson, J. C., Haapakangas, E.-L., Huhta, A., Nieminen, L., & Ullakonoja, R. (2014). *The diagnosis of reading in a second or foreign language*. Routledge.

Alderson, J. C., Nieminen, L., & Huhta, A. (2016). Characteristics of weak and strong readers in a foreign language. *The Modern Language Journal*, *100*(4), 853–879. https://doi.org/10.1111/modl.12367

Alharbi, N. S. M. (2017). *An investigation into the academic writing: Difficulties of Saudi postgraduate students* [Doctoral dissertation, University of Exeter]. http://hdl.handle.net/10871/33113

Ambjörnsdóttir, B., & Ingvarsdóttir, H. (2018). *Language development across the life span: The impact of English on education and work in Iceland* (Vol. 34). Springer. https://doi.org/10.1007/978-3-319-67804-7

Aryadoust, V., & Baghaei, P. (2016). Does EFL readers' lexical and grammatical knowledge predict their reading ability? Insights from a perceptron artificial neural network study. *Educational Assessment*, *21*(2), 135–156. https://doi.org/10.1080/10627197.2016.1166343

Austin, P. C., & Brunner, L. J. (2003). Type I error inflation in the presence of a ceiling effect. *The American Statistician*, *57*(2), 97–104. https://doi.org/10.1198/0003130031450

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Bardel, C. (2016). The lexicon of advanced L2 learners. In K. Hyltenstam (Ed.), *Advanced proficiency and exceptional ability in second languages (Vol. 51)* (pp. 73–109). De Gruyter Mouton.

Bardel, C., Erickson, G., & Österberg, R. (2019). Learning, teaching and assessment of second foreign languages in Swedish lower secondary school – Dilemmas and prospects. *Apples: Journal of Applied Language Studies*, *13*(1), 7–26. https://doi.org/10.17011/apples/urn.201903011687

Bartolotti J. & Marian V. (2017). Bilinguals' existing languages benefit vocabulary learning in a third language. *Language Learning*, *67*(1), 110–140. https://doi.org/10.1111/lang.12200

Bauer, L., & Nation, I. S. P. (1993). Word families. *International Journal of Lexicography*, *6*(4), 253–279. https://doi.org/10.1093/ijl/6.4.253

Bergström, D., Norberg, C., & Nordlund, M. (2022a). Do textbooks support incidental vocabulary learning? – a corpus-based study of Swedish intermediate EFL materials. *Education Inquiry*, Advance Online Publication, December 28, 2022. https://doi.org/10.1080/20004508.2022.2163050

Bergström, D., Norberg, C., & Nordlund, M. (2022b). "Words are picked up along the way" – Swedish EFL teachers' conceptualizations of vocabulary knowledge and learning. *Language Awareness*, *31*(4), 393–409. https://doi.org/10.1080/09658416.2021.1893326

Bergström, D., Norberg, C., & Nordlund, M. (2023). "The text comes first" – Principles guiding EFL materials developers' vocabulary content decisions. *Scandinavian Journal of Educational Research*, *67*(1), 154–168. https://doi.org/10.1080/00313831.2021.1990122

Biggs, J. (2003). Aligning teaching and assessment to curriculum objectives. *International Conference on Teaching and Learning in Higher Education: New trends and innovations.*

Bogaards, P. (2000). Testing L2 vocabulary knowledge at a high level: The case of the Euralex French tests. *Applied Linguistics*, *21*(4), 490–516. https://doi.org/10.1093/applin/21.4.490

Bolton, K., & Kuteeva, M. (2012). English as an academic language at a Swedish university: Parallel language use and the 'threat' of English. *Journal of Multilingual and Multicultural Development*, *33*(5), 429–447. https://doi.org/10.1080/01434632.2012.670241

Bonnet, A. & Siemund, P. (2018). *Foreign language education in multilingual classrooms*. John Benjamins.

Bonnet, G. (2004). *The assessment of pupils' skills in English in eight European countries 2002: A European Project*. European Network of Policy Makers for the Evaluation of Education Systems.

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

Brown, D., Stoeckel, T., Mclean, S., & Stewart, J. (2022). The most appropriate lexical unit for L2 vocabulary research and pedagogy: A brief review of the evidence. *Applied Linguistics*, *43*(3), 596–602. https://doi.org/10.1093/applin/amaa061

Brown, H. D. (2007). *Teaching by principles: An interactive approach to language pedagogy*. Pearson Education.

Brown, J., & Hudson, T. (2002). *Criterion-referenced language testing*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524803

Browne, C., Culligan, B., & Phillips, J. (2013a). *A new academic word list: The most important words for understanding academic text*. Retrieved December 9, 2022, from http://www.newgeneralservicelist.org/nawl-new-academic-word-list

Browne, C., Culligan, B., & Phillips, J. (2013b). *The new general service list*. Retrieved December 9, 2022, from http://www.newgeneralservicelist.org/

Busby, N. L. (2018). Comparing first and second language reading: The use of metacognitive strategies among Norwegian students. *Acta Didactica Norden*, *12*(2), 4–27. https://doi.org/10.5617/adno.5579

Busby, N. L. (2020a). *Presumptions, proficiencies, and parallel languages: Investigating academic English reading among Norwegian university students* [Doctoral dissertation, Norwegian University of Technology and Science, 2020:353].

Busby, N. L. (2020b). Words from where? Predictors of L2 English vocabulary among Norwegian university students. *ITL – International Journal of Applied Linguistics*. https://doi.org/10.1075/itl.19018.bus

Butler, Y. G. (2011). The implementation of communicative and task-based language teaching in the Asia-Pacific region. *Annual Review of Applied Linguistics*, *31*, 36–57. https://doi.org/10.1017/s0267190511000122

Bylund, J. (2022). *Everyday language practices and the interplay of ideologies, investment and identities – Language use and dispositions among young adolescents in multilingual urban setting in Sweden* [Doctoral dissertation, University of Gothenburg]. Acta universitatis gothoburgensis. Gothenburg studies in educational sciences 472. https://hdl.handle.net/2077/73698

Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, *130*(3), 278–288. https://doi.org/10.1016/j.cognition.2013.11.015

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*(1), 1–47. https://doi.org/10.1093/applin/i.1.1

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, *48*, 1-29. https://doi.org/10.18637/jss.v048.i06

Charles, M., & Pecorari, D. (2015). *Introducing English for academic purposes*. Routledge. https://doi.org/10.4324/9781315682129

Chen C., & Truscott J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics*, *31*(5), 693–713. https://doi-org.ezproxy.ub.gu.se/10.1093/applin/amq031

Chung, T. M., & Nation, I. S. P. (2003). Technical vocabulary in specialised texts. *Reading in a Foreign Language*, *15*(2), 103–116.

Chung, T. M., & Nation, I. S. P. (2004). Identifying technical vocabulary. *System*, *32*(2), 251–263. https://doi.org/10.1016/j.system.2003.11.008

Coady, J., & Nation, I. S. P. (1988). Vocabulary and reading. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 97–110). Longman.

Cobb, T. (2022). *Compleat Web VP v.2.6. [Vocabulary Profiler computer program]. Accessed 1 November 2022 at Lextutor homepage.* Retrieved December 9, 2022, from https://www.lextutor.ca/vp/comp/

Coe, R., Rauch, C. J., Kime, S., & Singleton, D. (2019). *Great teaching toolkit: Evidence review*. Evidence Based Education. https://assets.website-files.com/5ee28729f7b4a5fa99bef2b3/5ee9f507021911ae35ac6c4d_EBE_GTT_EVIDENCE%20REVIEW_DIGITAL.pdf

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.

Cowan, J. R. (1974). Lexical and syntactic research for the design of EFL reading materials. *TESOL Quarterly*, *8*(4), 389–399. http://dx.doi.org/10.2307/3585470

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*(2), 213–238. https://doi.org/10.2307/3587951

Coxhead, A. (2016). Acquiring academic and disciplinary vocabulary. In K. Hyland & P. Shaw (Eds.), *The Routledge handbook of English for academic purposes* (pp. 177–190). Routledge. https://doi.org/10.4324/9781315657455

Coxhead, A., & Boutorwick, T. (2018). Longitudinal vocabulary development in an EMI international school context: Learners and texts in EAL, maths, and science. *TESOL Quarterly*, *52*(3), 588–610. https://doi.org/10.1002/tesq.450

Cronbach, L. J. (1942). An analysis of techniques for diagnostic vocabulary testing. *The Journal of Educational Research*, *36*(3), 206–217. https://doi.org/10.1080/00220671.1942.10881160

Crooks, T. J., Kane, M. T., & Cohen, A. S. (1996). Threats to the valid use of assessments. *Assessment in Education: Principles, Policy & Practice*, *3*(3), 265–286. https://doi.org/10.1080/0969594960030302

Csöregh, A. M. (2022). *Fairest of them all?: Assessment identity development among Swedish student and novice teachers of English as a foreign language* (Doctoral dissertation, Stockholm University). DIVA.

Daller, M. H., & Phelan, D. (2013). Predicting international student study success. *Applied Linguistics Review*, *4*(1), 173–193. https://doi.org/10.1515/applirev-2013-0008

Dearden, J. (2015). *English as a medium of instruction – A growing global phenomenon*. British Council. https://doi.org/10.13140/RG.2.2.12079.94888

Delgadillo-Collazos, Y. (2020). Attitudes of students from bachelor's degree programs related to the implementation of the extensive reading technique in the English courses as a foreign language. *Revista Espacios*, *41*(44), 217–226. https://doi.org/10.48082/espacios-a20v41n44p16

Durrant, P. (2014). Discipline and Level Specificity in University Students' Written Vocabulary. *Applied Linguistics*, *35*(3), 328–356, https://doi.org/10.1093/applin/amt016

Durrant, P. (2016). To what extent is the Academic Vocabulary List relevant to university student writing? *English for Specific Purposes*, *43*, 49–61. https://doi.org/10.1016/j.esp.2016.01.004

Edgarsson, G. (2018). Academic vocabulary proficiency and reading comprehension among Icelandic secondary school students. In B. Ambjörnsdóttir & H. Ingvarsdóttir (Eds.), *Language development across the life span* (pp. 95–112). Springer. https://doi.org/10.1007/978-3-319-67804-7

Education First. (2019). *EF English proficiency index: A ranking of 100 countries and regions by English skills*.

Education First. (2021). *EF EPI. EF English proficiency index: A ranking of 112 countries and regions by English skills*. https://www.ef.com/assetscdn/WIBIwq6RdJvcD9bc8RMd/cefcom-epi-site/reports/2021/ef-epi-2021-english.pdf

Ellis, R. (1999). Factors in the incidental acquisition of second language vocabulary from oral input. In R. Ellis (Ed.), *Learning a second language through interaction* (pp. 35–61). John Benjamins Publishing Company. https://doi.org/10.1075/sibil.17.06ell

Elman, J. L. (2011). Lexical knowledge without a lexicon? *The Mental Lexicon*, *6*(1), 1–33. https://doi.org/10.1075/ml.6.1.01elm

Embretson, S. E., & Reise, S. P. (2013). *Item response theory*. Psychology Press.

Enever, J. (2018). *Policy and politics in global primary English*. Oxford University Press.

Enström, I. (2004). Ordförråd och ordinlärning – med särskilt fokus på avancerade inlärare. In K. Hyltenstam & I. Lindberg (Eds.), *Svenska som andraspråk – i forskning, undervisning och samhälle* (pp. 171–195). Studentlitteratur.

Erickson, G. (2004). *Engelska i åtta europeiska länder* [English in eight European countries]. Skolverket.

Erickson, G. (2010). Good practice in language testing and assessment: A matter of responsibility and respect. In T. Kao & Y. Lin (Eds.), *A new look at teaching and testing: English as subject and vehicle* (pp. 237–258). LTTC International Conference on English Language Teaching and Testing.

Erickson, G., Borger, L., & Olsson, E. (2022). National assessment of foreign languages in Sweden: A multifaceted and collaborative venture. *Language Testing*, *39*(3), 474–493. https://doi.org/10.1177/02655322221075067

Eriksson, L. (2021, September 6–10). *"It's a difficult transition" - Academic Second Language Reading in English Amongst Swedish Upper Secondary School Pupils and University Students* [Paper presentation]. ECER 2021 – European Conference on Educational Research, Geneva, Switzerland.

Eriksson, L. (2022). Difficulties in academic reading for EFL students: An initial investigation. *Language Teaching*, 1–4. https://doi.org/10.1017/S0261444822000246

Eriksson, L., Busby, N. & Warnby, M. (2023, January 11-12). *Preparing students for academic reading in English: A comparative study of Swedish and Norwegian university-preparatory curricula* [Paper presentation]. ISCOD - International Symposium on Comparative Didactics, Örebro University, Örebro, Sweden.

European Commission. (2012). *Final Report for the First European Survey on Language Competence*. Publications Office of the European Union.

European Commission. (2018). *Eurostat. 65% know at least one foreign language in the EU.* Retrieved December 9, 2022, from https://ec.europa.eu/eurostat/web/products-eurostat-news/-/EDN-20180926-1

Falk, Y., Lindqvist, C., & Bardel, C. (2015). The role of L1 explicit metalinguistic knowledge in L3 oral production at the initial state. *Bilingualism: Language and Cognition*, *18*(2), 227–235. https://doi.org/10.1017/s1366728913000552

Farrell, P. (1990). Vocabulary in ESP: A lexical analysis of the English of electronics and a study of semi-technical vocabulary. *CLCS Occasional Paper* No. 25. Trinity College.

Feast, V. (2002). The impact of IELTS scores on performance at university. *International Education Journal*, *3*(4), 70–85.

Feuer, M. J., Holland, P. W., Green, B. F., Bertenthal, M. W., & Hemphill, F. C. (1999). *Uncommon measures: Equivalence and linkage among educational tests*. ERIC.

Finndahl, I. (in progress). *Young students' language choice in Swedish compulsory school – Expectations, learning and assessment*. [Doctoral dissertation, University of Gothenburg].

Freimuth, H. (2020). Are academic English words learned incidentally? A Canadian case study. *BC TEAL Journal*, *5*(1), 32–43. https://doi.org/10.14288/bctj.v5i1.344

Fuster, C. (2022). *Lexical transfer in pedagogical translanguaging: Exploring intentionality in multilingual learners of Spanish* [Doctoral dissertation, Stockholm University]. DIVA.

Fuyun, S. (2002). A comparative study on the productive and academic vocabulary knowledge of Japanese and Chinese university students – Focusing on the students' performances in the two vocabulary tests. *Journal of International Development and Cooperation*, 113.

Gardner, D., & Davies, M. (2014). A New Academic Vocabulary List. *Applied Linguistics*, *35*(3), 305–327. https://doi.org/10.1093/applin/amt015

Gholaminejad, R., & Sarab, M. R. A. (2021). A comparison of the Academic Word List and the Academic Vocabulary List: Should the AVL replace the AWL? *TEFLIN Journal*, *32*(1), 167–182. https://doi.org/10.15639/teflinjournal.v32i1/167-182

Gipps, C. (1999). Socio-cultural aspects of assessment. *Review of Research in Education*, 24(1), 355–392. https://doi.org/10.2307/1167274

González-Fernández, B. (2022). Conceptualizing L2 vocabulary knowledge: An empirical examination of the dimensionality of word knowledge. *Studies in Second Language Acquisition*, *44*(4), 1124–1154. https://doi.org/10.1017/S0272263121000930

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6–10. https://doi.org/10.1177/074193258600700104

Gough, P. B., Hoover, W. A., & Peterson, C. L. (1996). Some observations on a Simple view of reading. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 1–13). Lawrence Erlbaum Associates Publishers.

Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly*, *25*(3), 375–406. https://doi.org/10.2307/3586977

Grabe, W. (2009). *Reading in a second language: moving from theory to practice*. Cambridge University Press.

Grabe, W., & Stoller, F. L. (1997). Reading and vocabulary development. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 98–122). Cambridge University Press. https://doi.org/10.1017/CBO9781139524643.009

Grabe, W., & Stoller, F. L. (2019). *Teaching and researching reading*. *Third edition*. Routledge.

Graves, K. (2016). Method, methods, and methodology: Historical trends and current debates. In G. Hall (Ed.), *The Routledge handbook of English teaching* (pp. 79–94). Routledge.

Graves, K., & Garton, S. (2017). An analysis of three curriculum approaches to teaching English in public-sector schools. *Language Teaching*, *50*(4), 441–482. https://doi.org/10.1017/S0261444817000155

Green, C. (2022). Computing curriculum time and input for the incidental learning of academic vocabulary. *Language Learning and Technology*, *26*(1), 1–21. https://doi.org/10125/73471

Greene, J. W., & Coxhead, A. (2015). *Academic vocabulary for middle school students: Research-based lists and strategies for key content areas*. Brookes Publishing.

Gyllstad, H. (2007). *Testing English collocations: Developing receptive tests for use with advanced Swedish learners* [Doctoral dissertation, Lund University]. Lund University Publications.

Gyllstad, H. (2013). Looking at L2 vocabulary knowledge dimensions from an assessment perspective – challenges and potential solutions. In C. Bardel, B. Laufer, & C. Lindqvist (Eds.), *L2 vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis* (pp. 11–28). EuroSLA – the European Second Language Association.

Gyllstad, H., McLean, S., & Stewart, J. (2020). Using confidence intervals to determine adequate item sample sizes for vocabulary tests: An essential but overlooked practice. *Language Testing*, 38(4), 558–579. https://doi.org/10.1177/0265532220979562

Gyllstad, H., Vilkaitė, L., & Schmitt, N. (2015). Assessing vocabulary size through multiple-choice formats: Issues with guessing and sampling rates. *ITL – International Journal of Applied Linguistics*, *166*(2), 278–306. https://doi.org/10.1075/itl.166.2.04gyl

Hambleton, R. K., & Swaminathan, H. (2013). *Item response theory: Principles and applications*. Springer Science & Business Media.

Hartshorn, K. J., & Hart, J. M. (2016). Comparing the academic word list with the academic vocabulary list: Analyses of frequency and performance of English language learners. *Journal of Language Teaching and Learning*, *6*(2), 70–87.

Hellekjær, G. O. (2005). The acid test: Does upper secondary EFL instruction effectively prepare Norwegian students for the reading of English textbooks at colleges and universities?

Hellekjær, G. O. (2009). Academic English reading proficiency at the university level: A Norwegian case study. *Reading in a Foreign Language*, *21*(2), 198–222.

Henriksen, B., & Danelund, L. (2015). Studies of Danish l2 learners' vocabulary knowledge and the lexical richness of their written production in English. In P. Pietilä, K. Doró & R. Pípalová (Eds.), *Lexical issues in L2 writing* (pp. 29–56). Cambridge Scholars Publishing.

Henry, A. (2012). *L3 Motivation* [Doctoral dissertation, University of Gothenburg]. University of Gothenburg.

Henry, A. (2014). Swedish students' beliefs about learning English in and outside of school. In D. Lasagabaster, A. Doiz, & J. M. Sierra (Eds.), *Motivation and foreign language learning: From theory to practice* (pp. 93–116). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.40.05hen

Hoffmann, R., Lagerkvist, C. J., Hagberg Gustavsson, M., & Holst, B. S. (2018). An empirical examination of the conceptualization of companion animals. *BMC Psychology*, *6*(15), 1–13. https://doi.org/10.1186/s40359-018-0228-1

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, *2*(2), 127–160. https://doi.org/10.1007/BF00401799

Hu, M. & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, *13*(1), 403–430. https://doi.org/10.26686/wgtn.12560354

Hulme, C., & Snowling, M. J. (2011). Children's reading comprehension difficulties: Nature, causes, and treatments. *Current Directions in Psychological Science*, *20*(3), 139–142. https://doi.org/10.1177/0963721411408673

Hult, F. M. (2012). English as a transcultural language in Swedish policy and practice. *TESOL Quarterly*, *46*(2), 230–257. https://doi.org/10.1002/tesq.19

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, *27*(1), 4–21. https://doi.org/10.1016/j.esp.2007.06.001

Hyland, K., & Tse, P. (2007). Is there an "academic vocabulary"? *Tesol Quarterly*, *41*(2), 235–253. https://doi.org/10.1002/j.1545-7249.2007.tb00058.x

Hyltenstam, K. (2004). *Engelskan, skolans språkundervisning och svensk språkpolitik*. Svenska språknämnden.

IELTS. (2022). *Setting IELTS entry scores*. https://www.ielts.org/for-organisations/setting-ielts-entry-scores

Institutet för språk och folkminnen. (2021). *Vilka är de största minoritetsspråken i Sverige?* https://www.isof.se

Kamimoto, T. (2008). *Nation's vocabulary levels test and its successors: A re-appraisal*. ProQuest Dissertations Publishing.

Kane, M. (2006). Content-related validity evidence in test development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 131–153). Lawrence Erlbaum Associates Publishers.

Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices*. Springer Science & Business Media.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, *73*(4), 440–464.

Krashen, S. (1992). The input hypothesis: An update. In J. Alatis (Ed.), *Georgetown University Round Table on Languages and Linguistics* (pp. 409–431). Georgetown University Press.

Krashen, S. (2003). *Explorations in language acquisition and use*. Heinemann.

Kremmel, B. (2016). Word families and frequency bands in vocabulary tests: Challenging conventions. *TESOL Quarterly*, *50*(4), 976–987. https://doi.org/10.1002/tesq.329

Kremmel, B., & Schmitt, N. (2018). Vocabulary levels test. In J. I. Liontas (Ed.), *The TESOL Encyclopedia of English Language Teaching* (pp. 1–7). John Wiley & sons. https://doi.org/10.1002/9781118784235.eelt0499

Kuppens, A. H. (2010). Incidental foreign language acquisition from media exposure. *Learning, Media and Technology*, *35*(1), 65–85. https://doi.org/10.1080/17439880903561876

Kuteeva, M. (2014). The parallel language use of Swedish and English: the question of 'nativeness' in university policies and practices. *Journal of Multilingual and Multicultural Development*, *35*(4), 332–344. https://doi.org/10.1080/01434632.2013.874432

Laufer, B. (1989). What percentage of text-lexis is essential for comprehension. In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316–323). Multilingual Matters.

Laufer, B. (1991). Knowing a word: What is so difficult about it? *English Teachers' Journal*, *42*(1), 82–86.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud & H. Béjoin (Eds.), *Vocabulary and applied linguistics* (pp. 126–132). Macmillan Academic and Professional. https://doi.org/10.1007/978-1-349-12396-4_12

Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics*, *19*(2), 255–271.

Laufer, B. (2021). Lemmas, flemma, word families, and common sense. *Studies in Second Language Acquisition*, *43*(5), 965–968. https://doi.org/10.1017/S0272263121000656

Laufer, B., & Aviad-Levitzky, T. (2017). What type of vocabulary knowledge predicts reading comprehension: Word meaning recall or word meaning recognition? *The Modern Language Journal*, *101*(4), 729–741. https://doi.org/10.1111/modl.12431

Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics*, *29*(4), 694–716. https://doi.org/https://doi-org.ezproxy.ub.gu.se/10.1093/applin/amn018

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, *54*(3), 399–436. https://doi.org/10.1111/j.0023-8333.2004.00260.x

Laufer, B., & Nation, I. S. P. (1995). Vocabulary Size and Use: Lexical Richness in L2 Written Production. *Applied Linguistics*, *16*(3), 307–322. https://doi.org/10.1093/applin/16.3.307

Laufer, B., & Nation, I. S. P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, *16*(1), 33–51. https://doi.org/10.1177/026553229901600103

Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, *48*(3), 365–391. https://doi.org/10.1111/0023-8333.00046

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical Threshold Revisited: Lexical Text Coverage, Learners' Vocabulary Size and Reading Comprehension. *Reading in a Foreign Language*, *22*(1), 15–30.

Laufer, B., & Rozovski-Roitblat, B. (2015). Retention of new words: Quantity of encounters, quality of task, and degree of knowledge. *Language Teaching Research*, 19(6), 687–711. htttps://doi.org/10.1177/13621688145597

Laufer, B., & Sim, D. D. (1985). Measuring and explaining the reading threshold needed for English for academic purposes texts. *Foreign Language Annals*, *18*(5), 405–411.

Leeming, P. (2014). Analysis of the vocabulary size test. *The Language Education Journal*(5), 73–87.

Leese, M. (2010). Bridging the gap: Supporting student transitions into higher education. *Journal of Further and Higher Education*, *34*(2), 239–251. https://doi.org/10.1080/03098771003695494

Leontjev, D., Huhta, A., & Tolvanen, A. (2022). L2 English vocabulary breadth and knowledge of derivational morphology: One or two constructs? *Language Testing*, *0*(0), 02655322221114015. https://doi.org/10.1177/026553222211140

Levelt, W. J. (1989). *Speaking: From intention to articulation*. The MIT Press.

Lewis, M. (1993). *The lexical approach: The state of English language teaching and a way forward*. Language Teaching Publications.

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*(10), 447–454. https://doi.org/10.1016/j.tics.2006.08.007

Lim Falk, M., & Holmberg, P. (2016). Paths to academic writing in a globalized world: A longitudinal study of content and language integrated learning in upper secondary school in Sweden. In S. Plane, C. Bazerman, F. Rondelli, C. Donahue, A. N. Applebee, C. Boré, P. Carline, M. Marquilló Larruy, P. Rogers & D. Russell (Eds.), *Recherche en écritures: Regards pluriels* (pp. 187–211). Université de Lorraine.

Lindberg, I. (2006). Bedömning av skolrelaterat ordförråd. In M. Olofsson (Ed.), *Symposium 2006: Bedömning, flerspråkighet och lärande* (pp. 83–107). HLS förlag.

Lindberg, I., & Johansson, S. (2019). The development of Swedish receptive vocabulary in CLIL: A multilingual perspective. In L.K. Sylvén, (Ed.), *Investigating content and language integrated learning: Insights from Swedish high schools* (pp. 236–258). Multilingual Matters.

Lissitz, R. W. (2009). *The concept of validity: Revisions, new directions and applications*. IAP.

Macaro, E., Curle, S., Pun, J., An, J., & Dearden, J. (2018). A systematic review of English medium instruction in higher education. *Language Teaching*, *51*(1), 36–76. https://doi.org/10.1017/S0261444817000350

Malmström, H., & Pecorari, D. (2022). *Svenskans och engelskans roll inom forskning och högre utbildning Språkval och internationalisering [Language choice and internationalisation: The roles of Swedish and English in research and higher education]*. Institutet för språk och folkminnen.

Malmström, H., Pecorari, D., & Shaw, P. (2018). Words for what? Contrasting university students' receptive and productive academic vocabulary needs. *English for Specific Purposes*, *50*, 28–39. https://doi.org/10.1016/j.esp.2017.11.002

Masrai, A. (2019). Vocabulary and reading comprehension revisited: Evidence for high-, mid-, and low-frequency vocabulary knowledge. *Sage Open*, *9*(2), 1–13. https://doi.org/10.1177/2158244019845182

Masrai, A., & Milton, J. (2017). Recognition vocabulary knowledge and intelligence as predictors of academic achievement in EFL context. *TESOL International Journal*, *12*(1), 128–142.

Masrai, A., & Milton, J. (2018). Measuring the contribution of academic and general vocabulary knowledge to learners' academic achievement. *Journal of English for Academic Purposes*, *31*, 44–57. https://doi.org/10.1016/j.jeap.2017.12.006

Mauranen, A. (2012). *Exploring ELF: Academic English shaped by non-native speakers*. Cambridge University Press.

Mauranen, A., Hynninen, N., & Ranta, E. (2016). English as the academic lingua franca. In K. Hyland & P. Shaw (Eds.), *The Routledge handbook of English for academic purposes* (pp. 44–55). Routledge. https://doi.org/10.4324/9781315657455

McLean, S. (2018). Evidence for the adoption of the flemma as an appropriate word counting unit. *Applied Linguistics*, *39*(6), 823–845. https://doi.org/10.1093/applin/amx003

McLean, S. (2021). The coverage comprehension model, its importance to pedagogy and research, and threats to the validity with which it is operationalized. *Reading in a Foreign Language*, *33*(1), 126–140.

McLean, S., & Kramer, B. (2015). The creation of a new vocabulary levels test. *Shiken*, *19*(2), 1–11.

McLean, S., Kramer, B., & Stewart, J. (2015). An empirical examination of the effect of guessing on vocabulary size test scores. *Vocabulary Learning and Instruction*, *4*(1), 26–35. https://doi.org/10.7820/vli.v04.1.mclean.et.al

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (pp. 13–103). Macmillan Publishing.

Messick, S. (1996). Validity and washback in language testing. *Language Testing*, *13*(3), 241–256.

Milton, J. (2009). *Measuring second language vocabulary acquisition*. Multilingual Matters.

Milton, J. (2022). Vocabulary denial and the false god of structuralism in Ofsted's 2021 Curriculum research review for languages. *The Language Learning Journal*, *50*(2), 156–171. https://doi.org/10.1080/09571736.2022.2045680

Milton, J., & Hopwood, O. (2022). *Vocabulary in the foreign language curriculum: Principles for effective instruction*. Taylor & Francis.

Milton, J., & Treffers-Daller, J. (2013). Vocabulary size revisited: the link between vocabulary size and academic achievement. *Applied Linguistics Review*, *4*(1), 151–172. https://doi.org/10.1515/applirev-2013-0007

Montgomery, S. L. (2013). *Does science need a global language?: English and the future of research*. The University of Chicago Press.

Myrberg, E., & Rosén, M. (2009). Direct and indirect effects of parents' education on reading achievement among third graders in Sweden. *British Journal of Educational Psychology*, *79*(4), 695–711. https://doi.org/10.1348/000709909X453031

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, *47*(1), 91–108. https://doi.org/10.1002/RRQ.011

Nagy, W., Herman, P., & Anderson, R. (1985). Learning words from context. *Reading Research Quarterly*, 233–253.

Nation, I. S. P. (1983). Testing and teaching vocabulary. *Guidelines*, *5*(1), 12–25.

Nation, I. S. P. (1996). The four strands of a language course. *TESOL in Context*, *6*(1), 7–12. https://doi.org/10.26686/wgtn.12560387.v1

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge : Cambridge University Press.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, *63*(1), 59–82. https://doi.org/10.3138/cmlr.63.1.59

Nation, I. S. P. (2007). The four strands. *International journal of innovation in language learning and teaching*, *1(*(2), 2–13. https://doi.org/10.2167/illt039.0

Nation, I. S. P. (2008). *Teaching vocabulary: Strategies and techniques*. Heinle Cengage Learning.

Nation, I. S. P. (2013). *Learning vocabulary in another language. Second Edition.* Cambridge University Press.

Nation, I. S. P. (2014). How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language*, *26*(2), 1–16. https://doi.org/10125/66881

Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, *31*(7), 9–13. https://doi.org/http://www.jalt-publications.org/archive/tlt/2007/07_2007TLT.pdf

Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6–19). Cambridge University Press.

Nation, I. S. P., & Yamamoto, A. (2012). Applying the four strands to language learning. *International Journal of Innovation in English Language Teaching IP*.

Newton, J. (2020). Approaches to learning vocabulary inside the classroom. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 255–270). Routledge. https://doi.org/10.4324/9780429291586-17

Nordic Council of Ministers. (2007). *Declaration on a Nordic language policy*. Nordic Council.

O'Sullivan, B., & Weir, C. J. (2011). Test development and validation. In B. O'Sullivan (Ed.), *Language testing: Theories and practices* (pp. 13–32). Palgrave MacMillan.

Ohlsson, E. (2021). *Den synliggjorda vokabulären och praktiken: Gymnasieelevers akademiska skrivande på svenska* (Doctoral dissertation, University of Gothenburg).

Olmos, C. (2009). An assessment of the vocabulary knowledge of students in the final year of secondary education. Is their vocabulary extensive enough? *International Journal of English Studies*, *9*(3), 73–90.

Olsson, E. (2015). Progress in English academic vocabulary use in writing among CLIL and non-CLIL students in Sweden. *Moderna Språk*, *109*(2), 51–74.

Olsson, E. (2016). *On the impact of extramural English and CLIL on productive vocabulary* [Doctoral dissertation, University of Gothenburg].

Olsson, E., & Sylvén, L. K. (2015). Extramural English and academic vocabulary. A longitudinal study of CLIL and non-CLIL students in Sweden. *Apples: Journal of Applied Language Studies*, *9*(2), 77–103. https://doi.org/10.17011/apples/urn.201512234129

Palm, C. (2023). Diskurser om skrivande i styrdokument för svenska som andraspråk på grundläggande och gymnasial nivå inom vuxenutbildning. *Nordic Journal of Literacy Research*, *9*(1), 1–18. http://dx.doi.org/10.23865/njlr.v9.3529

Paquot, M. (2010). Academic vocabulary in learner writing: From extraction to analysis. Bloomsbury Publishing.

Paribakht, T. S., & Wesche, M. B. (1993). Reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal*, *11*(1), 9–29. https://doi.org/10.18806/tesl.v11i1.623

Paribakht, T. S., & Wesche, M. B. (1999). Reading and "incidental" L2 vocabulary acquisition: An introspective study of lexical inferencing. *Studies in Second Language Acquisition*, *21*(2), 195–224. https://doi.org/10.1017/s027226319900203x

Parkvall, M. (2016). *Sveriges språk i siffror : vilka språk talas och av hur många?* Språkrådet: Morfem.

Parkvall, M. (2019). *Den nya mångfalden*. Makadam förlag.

Pecorari, D., & Malmström, H. (2018). At the crossroads of TESOL and English medium instruction. *TESOL Quarterly*, *52*(3), 497–515. https://doi.org/10.1002/tesq.470

Pecorari, D., Shaw, P., & Malmström, H. (2019). Developing a new academic vocabulary test. *Journal of English for Academic Purposes*. https://doi.org/10.1016/j.jeap.2019.02.004

Pecorari, D., Shaw, P., Irvine, A., Malmström, H., & Mežek, Š. (2012). Reading in tertiary education: Undergraduate student practices and attitudes. *Quality in Higher Education*, *18*(2), 235–256. https://doi.org/10.1080/13538322.2012.706464

Pecorari, D., Shaw, P., Malmström, H., & Irvine, A. (2011). English textbooks in parallel-language tertiary education. *TESOL Quarterly*, *45*(2), 313–333. https://doi.org/10.5054/tq.2011.247709

Pellicer-Sánchez, A. (2017). Learning L2 collocations incidentally from reading. *Language Teaching Research*, *21*(3), 381–402. https://doi.org/10.1177/1362168815618421

Perander, K., Londen, M., Holm, G., & Tiihonen, S. (2020). Becoming a university student: An emotional rollercoaster. *Högre Utbildning*, *10*(1), 1–12. https://doi.org/10.23865/hu.v10.1462

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227–247). Blackwell Publishing. https://doi.org/10.1002/9780470757642.ch13

Peters, E. (2018). The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *ITL – International Journal of Applied Linguistics*, *169*(1), 142–168. https://doi.org/10.1075/itl.00010.pet

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, *40*(3), 551–577. https://doi.org/10.1017/s0272263117000407

Peters, E., Noreillie, E.S., Heylen, K., Bulté, B. & Desmet, P. (2019). The impact of instruction and out-of-school exposure to foreign language input on learners' vocabulary knowledge in two languages. *Language Learning, 69*(3), 747–782. https://doi.org/10.1111/lang.12351

Phillipson, R. (1992). *Linguistic imperialism*. Oxford University Press.

Pitts, M., White, H., & Krashen, S. (1989). Acquiring the second language vocabulary through reading: A replication of the Clockwork Orange study using second language acquirers. *Reading in a Foreign Language*, *5*(2), 271–275.

Praninskas, J. (1972). *American university word list*. Longman Group Limited.

Puimège, E., & Peters, E. (2019). Learners' English vocabulary knowledge prior to formal instruction: The role of learner-related and word-related variables. *Language Learning*, *69*(4), 943–977. https://doi.org/10.1111/lang.12364

Qi, H. (2016). A corpus-based comparison between the Academic Word List and the Academic Vocabulary List (Master of Arts, *Electronic Thesis and Dissertation Repository*, The University of Western Ontario, Paper 3556). http://ir.lib.uwo.ca/etd/3556

Qian, D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, *56*(2), 282–308. https://doi.org/10.3138/cmlr.56.2.282

Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52(3), 513–536. https://doi.org/10.1111/1467-9922.00193

Rea, L.M., & Parker, R.A. (2014). *Designing and conducting survey research: A comprehensive guide*. John Wiley & Sons.

Read, J. (2000). *Assessing vocabulary*. Cambridge University Press.

Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 209–227). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.10.15rea

Read, J., & Chapelle, C. A. (2001). A framework for second language vocabulary assessment. *Language Testing*, *18*(1), 1–32. https://doi.org/10.1177/0265532201018001

Read, J., & Dang, T. N. Y. (2022). Measuring depth of academic vocabulary knowledge. *Language Teaching Research*, Advance Online Publication, July 11, 2022. https://doi.org/10.1177/13621688221105913

Ribeck, J., Jansson, H., & Sköldberg, E. (2014). Från aspekt till övergripande–en ordlista över svensk akademisk vokabulär. *Nordiske Studier i Leksikografi* (12). https://tidsskrift.dk/nsil/article/view/21029

Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching*. Cambridge University Press.

Robinson, P. (1988, March 8–13). *Procedural and declarative knowledge in vocabulary learning: Communication and the language learner's lexicon* [paper presentation]. The Annual Meeting of the Teachers of English to Speakers of Other Languages (22nd, Chicago, IL). https://eric.ed.gov/?id=ED296585

RStudio Team. (2020). *RStudio: Integrated development for R*. RStudio. http://rstudio.com/

Salö, L. (2010). *Engelska eller svenska? En kartläggning av språksituationen inom högre utbildning och forskning*. Språkrådet.

Salö, L. (2016). *Languages and linguistic exchanges in Swedish academia practices, processes, and globalizing markets* [Doctoral dissertation, Stockholm University]. DIVA.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, *12*(3), 329–363. https://doi.org/10.1177/1362168808089921

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave MacMillan.

Schmitt, N. (2019). Understanding vocabulary acquisition, instruction, and assessment: A research agenda. *Language Teaching*, *52*(2), 261–274. https://doi.org/10.1017/S0261444819000053

Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, *47*(4), 484–503. https://doi.org/10.1017/S0261444812000018

Schmitt, N., Cobb, T., Horst, M., & Schmitt, D. (2017). How much vocabulary is needed to use English? Replication of van Zeeland & Schmitt (2012), Nation (2006) and Cobb (2007). *Language Teaching*, *50*(2), 212–226. https://doi.org/10.1017/S0261444815000075

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, *95*(1), 26–43. https://doi.org/ https://doi.org/10.1111/j.1540-4781.2011.01146.x

Schmitt, N., Nation, I. S. P., & Kremmel, B. (2020). Moving the field of vocabulary assessment forward: The need for more rigorous test development and validation. *Language Teaching*, *53*(1), 109–120. https://doi.org/10.1017/S0261444819000326

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, *18*(1), 55–88. https://doi.org/10.1177/026553220101800103

Serpell, J. A. (2009). Having our dogs and eating them too: Why animals are a social issue. *Journal of Social Issues*, *65*(3), 633–644. https://doi.org/10.1111/j.1540-4560.2009.01617.x

SFS 2003:460 [Swedish Code of Statuses 2003:460] (2003). *Lag om etikprövning av forskning som avser människor* (The Ethical Review Act). Sveriges Riksdag.

SFS 2009:600 [Swedish Code of Statuses 2009:600]. *Språklag* (The Language Act). Sveriges Riksdag.

SFS 2010:800 [Swedish Code of Statuses 2010:800]. *Skollagen* (The Education Act). Sveriges Riksdag.

SFS 2017:1128 [Swedish Code of Statuses 2017:1128]. (2017). *Diskrimineringslagen* (The Discrimination Act). Sveriges Riksdag.

Shaw, P., & McMillion, A. (2011). Components of success in academic reading tasks for Swedish students. *Ibérica* (22), 141–162. http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-65993

Siegel, J. (2019). Identifying priorities for Action Research in Swedish EFL classrooms. In B. Ljung Egeland, T. Roberts, E. Sandlund, & P. Sundqvist (Eds.), *Klassrumsforskning och språk(ande): Rapport från ASLA-symposiet i Karlstad, 12–13 april, 2018* (pp. 259–278) [Classroom research and language/languaging: Papers from the ASLA symposium in Karlstad, 12–13 April, 2018]. Karlstad University Press.

Siegel, J. (2022). *Teaching English in secondary school: A handbook for teachers*. Studentlitteratur.

Simensen, A. M. (2010). English in Scandinavia: a success story. In D. Wyse, R. Andrews, & J. Hoffman (Eds.), *The Routledge international handbook of English, language and literacy teaching* (pp. 472–483). Routledge.

Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, *31*(4), 487–512. https://doi.org/10.1093/applin/amp058

Singleton, D. (2000). *Language and the lexicon*. Oxford University Press. https://doi.org/10.4324/9781315824796

Skjelde, K. M., & Coxhead, A. (2020). Mind the gap: Academic vocabulary knowledge as a predictor of English grades. *Acta Didactica Norden*, *14*(3), 20 pages. https://doi.org/10.5617/adno.7975

Skolverket [National Agency for Education]. (2012b). Overview of the Swedish upper secondary school. Skolverket. Retrieved 25 November, 2022, from https://www.skolverket.se/download/18.6bfaca41169863e6a659a16/1553964183391/pdf2748.pdf

Skolverket [National Agency for Education]. (2014). *Vanliga begrepp vid översättning av betyg inom gymnasieskolan* (Common terminology when translating upper secondary school grades). Skolverket. Retrieved 25 November, 2022, from https://www.skolverket.se/download/18.189c87ae1623366ff375a3/1521546898821/Common-terminology-when-translating-upper-secondary-school-grades.pdf

Skolverket [National Agency for Education]. (2011). *Läroplan, examensmål och gymnasiegemensamma ämnen för gymnasieskola 2011*. Skolverket. http://www.skolverket.se/publikationer?id=2705

Skolverket [National Agency for Education]. (2012a). *Internationella språkstudien 2011 : Elevernas kunskaper i engelska och spanska*. Skolverket.

Skolverket [National Agency for Education]. (2013). *Curriculum for the upper-secondary school*. Skolverket. https://www.skolverket.se/getFile?file=2975

Skolverket [National Agency for Education]. (2015). *Undervisningstid och heltidsstudier i gymnasieskolan.* Skolverket. https://www.skolverket.se/download/18.6bfaca41169863e6a65b8e5/1553966209815/pdf3485.pdf

Skolverket [National Agency for Education]. (2017). *Ordförrådet – en framgångsfaktor*. Skolverket. https://larportalen.skolverket.se/LarportalenAPI/api-v2/document/path/larportalen/material/inriktningar/5-las-skriv/F%C3%B6rskoleklass/029-vardagssprak-amnessprak/del_05/Material/Flik/Del_05_MomentA/Artiklar/M29_05A_01_ordforrad.doc

Skolverket [National Agency for Education]. (2018). *Curriculum for the compulsory school, preschool class and school-age educare (revised 2018)*. Skolverket. https://www.skolverket.se/publikationer?id=3984

Skolverket. (2020). *Att planera, bedöma och ge återkoppling. Stöd för undervisning*. Skolverket. https://www.skolverket.se/download/18.2a6bcc30176f24c3889241/1611039323467/pdf7689.pdf

Skolverket [National Agency for Education]. (2021a). *Kommentarmaterial till ämnesplanerna i moderna språk och engelska*. Skolverket. https://www.skolverket.se/publikationer?id=7842

Skolverket [National Agency for Education]. (2021b). *Timplan för grundskolan* Skolverket. https://www.skolverket.se/undervisning/grundskolan/laroplan-och-kursplaner-for-grundskolan/timplan-for-grundskola

Skolverket [National Agency for Education]. (2022a). *Gymnasieprogrammen*. Skolverket. https://www.skolverket.se/undervisning/gymnasieskolan/laroplan-program-och-amnen-i-gymnasieskolan/gymnasieprogrammen

Skolverket [National Agency for Education]. (2022b). *Jämför ämnesplanerna. Ämnesplan i engelska på gymnasial nivå.* Skolverket. https://www.skolverket.se/download/18.3770ea921807432e6c72f3e/1657719361923/jamforelsedok_Engelska.pdf

Skolverket [National Agency for Education]. (2022c). *Kursplan för grundskolan – Engelska* (Läroplan (Lgr22) för grundskolan, förskoleklassen och fritidshemmet (SKOLFS 2010:37)). Skolverket. https://www.skolverket.se/undervisning/grundskolan/laroplan-och-kursplaner-for-grundskolan/laroplan-lgr22-for-grundskolan-samt-for-forskoleklassen-och-fritidshemmet?url=-996270488%2Fcompulsorycw%2Fjsp%2Fsubject.htm%3FsubjectCode%3DGRGRENG01%26tos%3Dgr&sv.url=12.5dfee44715d35a5cdfa219f

Skolverket [National Agency for Education]. (2022d). *Kursplan för grundskolan – Idrott och hälsa* (*Läroplan (Lgr22) för grundskolan, förskoleklassen och fritidshemmet (SKOLFS 2010:37)*). Skolverket. https://www.skolverket.se/undervisning/grundskolan/laroplan-och-kursplaner-for-grundskolan/laroplan-lgr22-for-grundskolan-samt-for-forskoleklassen-och-fritidshemmet?url=-996270488%2Fcompulsorycw%2Fjsp%2Fsubject.htm%3FsubjectCode%3DGRGRIDR01%26tos%3Dgr&sv.url=12.5dfee44715d35a5cdfa219f#anchor4

Skolverket [National Agency for Education]. (2022e). *Sveriges officiella statistik: Gymnasieskolan, betyg och studieresultat, läsåret 2020/2021*. Skolverket. https://www.skolverket.se/skolutveckling/statistik/sok-statistik-om-forskola-skola-och-vuxenutbildning?sok=SokC&verkform=Gymnasieskolan&omrade=Betyg%20och%20studieresultat&lasar=2020/21&run=1

Skolverket [National Agency for Education]. (2022f). *Ämnesbetygsreformen*. Skolverket. https://www.skolverket.se/undervisning/gymnasieskolan/aktuella-forandringar-pa-gymnasial-niva/amnesbetygsreformen

Skolverket [National Agency for Education]. (2022g). *Ämnesplaner i gymnasieskolan – Engelska* (*Läroplanen för gymnasieskolan (SKOLFS 2022:13)*). Skolverket. https://www.skolverket.se/undervisning/gymnasieskolan/laroplan-program-och-amnen-i-gymnasieskolan/gymnasieprogrammen/amne?url=1530314731%2Fsyllabuscw%2Fjsp%2Fsubject.htm%3FsubjectCode%3DENG%26lang%3Dsv%26tos%3Dgy%26p%3Dp&sv.url=12.5dfee44715d35a5cdfa92a3

Skolverket [National Agency for Education]. (2022h). *Ämnesplaner i gymnasieskolan – Samhällskunskap* (*Läroplanen för gymnasieskolan (SKOLFS 2022:13)*). Skolverket. https://www.skolverket.se/undervisning/gymnasieskolan/laroplan-program-och-amnen-i-gymnasieskolan/gymnasieprogrammen/amne?url=-996270488%2Fsyl-labuscw%2Fjsp%2Fsubject.htm%3FsubjectCode%3DSAM%26version%3D4%26tos%3Dgy&sv.url=12.5dfee44715d35a5cdfa92a3#anchor4:~:text=Samh%C3%A4llskunskap%20(pdf)

Skolverket [National Agency for Education]. (2023). *Läslyftet i skolan.* Skolverket. Retrieved 25 November, 2022, from https://www.skolverket.se/skolutveckling/kurser-och-utbildningar/laslyftet-i-skolan

Skoufaki, S., & Petrić, B. (2021). Exploring polysemy in the Academic Vocabulary List: A lexicographic approach. *Journal of English for Academic Purposes*, *54*, 101038. https://doi.org/10.1016/j.jeap.2021.101038Get

Snoder, S. (2022). *"Det är inte så att vi exklusivt pratar svenska.": Flerspråkighetens möjliga utrymmen i låg-och mellanstadieklassrum. En studie av flerspråkiga grundskollärares perspektiv* (Doctoral dissertation, Stockholm University).

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olsen, & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112-133). Cambridge University Press. https://doi.org/10.1017/cbo9780511609664.008

Snow, C. E., Lawrence, J. F., & White, C. (2009). Generating knowledge of academic language among urban middle school students. *Journal of Research on Educational Effectiveness*, *2*(4), 325–344. https://doi.org/10.1080/19345740903167042

SOU 2002:27 [Official Reports of the Swedish Government]. (2002). *Mål i mun – Förslag till handlingsprogram för svenska språket*. https://www.regeringen.se/rattsliga-dokument/statens-offentliga-utredningar/2002/04/sou-200227/

SOU 2018:3 [Official Reports of the Swedish Government]. (2018). *En strategisk agenda för internationalisering*.

Språkrådet. (2022). *Språken i Sverige*. https://www.isof.se/lar-dig-mer/kunskapsbanker/lar-dig-mer-om-flersprakighet/spraken-i-sverige

Staehr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *The Language Learning Journal*, *36*(2), 139–152. https://doi.org/10.1080/09571730802389975

Statens medieråd [Swedish Media Council]. (2019). *Ungar & medier 2019* [Kids & Media 2019].

Statistics Sweden. (2012). *Theme report 2012 – 2. Theme: Education. The period between upper secondary school and university/higher education*. Statistics Sweden. https://www.scb.se

Statistics Sweden. (2022). *Befolkningsstatistik*. Retrieved from https://www.scb.se/hitta-statistik/statistik-efter-amne/befolkning/befolkningens-sammansattning/befolkningsstatistik/#_

Stelma, J. (2009). What is communicative language teaching. In S. Hunston & D. Oakey (Eds.), *Introducing applied linguistics: Concepts and skills* (pp. 53–59). Routledge.

Stewart, J., Stoeckel, T., McLean, S., & Pinchbeck, G. G. (2021). What the research shows about written receptive vocabulary testing: A reply to Webb. *Studies in Second Language Acquisition*, *43*(2), 462–471. https://doi.org/10.1017/S0272263121000437

Stewart, J., & White, D. A. (2011). Estimating guessing effects on the vocabulary levels test for differing degrees of word knowledge. TESOL Quarterly, 45(2), 370–380. https://doi.org/10.5054/tq.2011.254523

Stoeckel, T., Ishii, T., & Bennett, P. (2020). Is the lemma more appropriate than the flemma as a word counting unit? *Applied Linguistics*, *41*(4), 601–606.

Stoeckel, T., McLean, S., & Nation, I. S. P. (2021). Limitations of size and levels tests of written receptive vocabulary knowledge. *Studies in Second Language Acquisition*, *43*(1), 181–203. https://doi.org/10.1017/S0272263121000437

Stuart, M., Stainthorp, R., & Snowling, M. (2008). Literacy as a complex activity: Deconstructing the simple view of reading. *Literacy*, *42*(2), 59–66. https://doi.org/10.1111/j.1741-4369.2008.00490.x

Sturk, E. (2022). *Writing across the curriculum in compulsory school in Sweden* (Doctoral dissertation, Umeå university).

Sundqvist, P. (2009). *Extramural English matters out-of-school English and its impact on Swedish* ninth graders' oral proficiency and vocabulary [Doctoral dissertation, Karlstads universitet].

Sundqvist, P. (2019). Sweden and informal language learning. In M. Dressman & R. Sadler (Eds.), *The handbook of informal language learning* (pp. 319–332). John Wiley & Sons.

Sundqvist, P., & Sylvén, L. K. (2016). *Extramural English in teaching and learning: from theory and research to practice*. Palgrave Macmillan.

Sundqvist, P., & Wikström, P. (2015). Out-of-school digital gameplay and in-school L2 English vocabulary outcomes. *System*, *51*(July 2015), 65–76. https://doi.org/10.1016/j.system.2015.04.001

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. *Input in Second Language Acquisition*, *15*, 165–179.

Swedish Council for Higher Education. (2016). *Can excellence be achieved in homogeneous student groups? A report on the governmental assignment to survey and analyse the work of Swedish higher education institutions on widening access and widening participation.* Swedish Council for Higher Education. Retrieved from www.uhr.se/english.

Swedish Council for Higher Education. (2017). *Tidiga avhopp från högskolan. Rapport 2017:17*. Retrieved from https://www.uka.se/download/18.2b48d4bc15ec792491a331e/1507896110834/rapport-2017-10-12-tidiga-avhopp-fran-hogskolan.pdf

Sylvén, L. K. (2006a, 10–11 November 2005). *Extramural exposure to English among Swedish school students* [Paper presentation]. ASLA Symposium, Växjö, Sweden.

Sylvén, L. K. (2006b). How is extramural exposure to English among Swedish school students used in the CLIL classroom. *VIEWS – Vienna English working papers*, *15*(3), 47–53.

Sylvén, L. K. (2019). *Investigating content and language integrated learning: Insights from Swedish high schools*. Multilingual Matters.

Sylven, L. K., & Ohlander, S. (2014). The CLISS Project: Receptive vocabulary in CLIL versus non-CLIL groups. *Moderna språk*, *108*(2), 80–114.

Sylvén, L. K., & Ohlander, S. (2019). English receptive vocabulary. In L. K. Sylvén (Ed.), *Investigating content and language integrated learning: Insights from Swedish high schools* (pp. 101–116). Multilingual Matters.

Sylvén, L. K., & Sundqvist, P. (2012). Gaming as extramural English L2 learning and L2 proficiency among young learners. *ReCALL: The Journal of EUROCALL*, *24*(3), 302–321. https://doi.org/10.1017/S095834401200016X

Therova, D. (2020). Review of academic word lists. *TESL-EJ*, *24*(1), 1–15. Retrieved November 4, 2020, from http://tesl-ej.org/pdf/ej93/a5.pdf

Thornbury, S. (2016). Communicative language teaching in theory and practice. In G. Hall (Ed.), *The Routledge handbook of English language teaching* (pp. 224–237). Routledge.

Tinkham, T. (1997). The effects of semantic and thematic clustering on the learning of second language vocabulary. *Second Language Research*, *13*(2), 138–163. https://doi.org/10.1191/026765897672376469

Townsend, D., Filippini, A., Collins, P., & Biancarosa, G. (2012). Evidence for the importance of academic word knowledge for the academic achievement of diverse middle school students. *The Elementary School Journal*, *112*(3), 497–518.

Trenkic, D., & Warmington, M. (2019). Language and literacy skills of home and international university students: How different are they, and does it matter? *Bilingualism: Language and Cognition*, *22*(2), 349–365. https://doi.org/10.1017/S136672891700075X

Tschirner, E. (2004). Breadth of vocabulary and advanced English study: An empirical investigation. *Electronic Journal of Foreign Language Teaching*, *1*(1), 27–39.

Valipouri, L., & Nassaji, H. (2013). A corpus-based study of academic vocabulary in chemistry research articles. *Journal of English for Academic Purposes*, *12*(4), 248–263. https://doi.org/10.1016/j.jeap.2013.07.001

van Zeeland, H. (2013). L2 vocabulary knowledge in and out of context: Is it the same for reading and listening? *Australian Review of Applied Linguistics*, *36*(1), 52–70.

Van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System*, *41*(3), 609–624. https://doi.org/10.1016/j.system.2013.07.012

Vetenskapsrådet [Swedish Research Council]. (2017). *God forskningssed*. *Reviderad utgåva*. Vetenskapsrådet.

Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics*, *24*(1), 56–89. https://doi.org/10.1093/applin/24.1.56

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, *61*(1), 219–258. https://doi.org/10.1111/j.1467-9922.2010.00593.x

Warnby, M. (2019a, August 28). *Exploring the Swedish and English passive academic vocabulary of upper secondary students* [Paper presentation]. EuroSLA29, Lund, Sweden.

Warnby, M. (2019b, July 1–3). *Investigating upper secondary students' Swedish and English receptive academic vocabulary* [Paper presentation]. Vocab@Leuven *2019*, Leuven, Belgium.

Warnby, M. (2021, August 15–20). *Do leisure activities promote upper secondary students' acquisition of academic vocabulary?* [Paper presentation]. AILA2021 World Congress of Applied Linguistics, (AILA2020 postponed), Groningen, Netherlands.

Warnby, M. (2022). Receptive academic vocabulary knowledge and extramural English involvement – is there a correlation? *ITL – International Journal of Applied Linguistics*, *173*(1), 120–152. https://doi.org/10.1075/itl.21021.war

Warnby, M. (2023). Academic vocabulary knowledge among adolescents in university preparatory programmes. *Journal of English for Academic Purposes*, *61*(2023) 101203, 1–14. https://doi.org/10.1016/j.jeap.2022.101203

Warnby, M., & Lemmouh, Z. (2021, October 1–3). *What? of academic lexis in curriculum learning outcome – The case of Sweden* [Paper presentation]. 3rd International Conference on Language Education and Research, Greece (Online).

Warnby, M., Malmström, H., & Bardel, C. (2022). Ett svenskt akademiskt vokabulärtest – Revidering, validering och förslag på vidareutveckling [A Swedish Academic Vocabulary Test – Revision, validation, and suggestions for further development]. *Nordic Journal of Literacy Research*, *8*(1), 106–127. https://doi.org/10.23865/njlr.v8.3282

Warnby, M., Malmström, H., & Yang Hansen, K. (2023). Linking scores from two English academic vocabulary tests – the VLT-Ac and the AVT. *Language Testing*. Advance Online Publication, January 12, 2023. https://doi.org/10.1177/02655322221145643

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46–65. https://doi.org/10.1093/applin/aml048

Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition*, *30*(1), 79–95. https://doi.org/10.1017/S0272263108080042

Webb, S. (2020). *Incidental vocabulary learning*. Routledge. https://doi.org/10.4324/9780429291586-15

Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, *63*(1), 91–120. https://doi.org/10.1111/j.1467-9922.2012.00729.x

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL – International Journal of Applied Linguistics*, *168*(1), 33–69. https://doi.org/10.1075/itl.168.1.02web

Wesche, M. B., & Paribakht, T. (1999). *Incidental L2 vocabulary acquisition: Theory, current research, and instructional implications*. Cambridge University Press.

West, M. (1953). *A general service list of English words*. Longman, Green and Co.

Wilkins, D. A. (1972). *Linguistics in language teaching* (Vol. 111). Edward Arnold London.

Wray, A. (2015). Why are we so sure we know what a word is? In J. R. Taylor (Ed.), *The Oxford handbook of the word* (pp. 725–750). Oxford University Press. https://doi-org.ezproxy.ub.gu.se/10.1093/oxfordhb/9780199641604.013.032

Young Academy of Sweden/Sveriges Unga Akademi (2022). *A beginner's guide to Swedish academia*. Young Academy of Sweden.

Xi, X., & Sawaki, Y. (2017). Methods of test validation. In E. Shohamy, I. G. Or & S. May (Eds.), *Language testing and assessment* (pp. 193–209). Springer. https://doi.org/10.1007/978-3-319-02261-1_14

Xue, G., & Nation, I. S. P. (1984). A university word list. *Language Learning and Communication*, *3*(2), 215–229.

Zhang, S., & Zhang, X. (2020). The relationship between vocabulary knowledge and L2 reading/listening comprehension: A meta-analysis. *Language teaching research*, *26*(4), p. 696-724. https://doi.org/10.1177/1362168820913998

Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. *Second Language Vocabulary Acquisition*, *6*(1), 5–19. https://doi.org/10.1017/cbo9781139524643.003

Appendices

Appendix A

Stockholms universitet

Förfrågan om deltagande i forskningsprojekt

Hej!

Jag heter Marcus Warnby och skulle behöva Er hjälp.

Jag är doktorand vid Stockholms universitet med en bakgrund som gymnasielärare. Mitt avhandlingsarbete är inriktat på svensk och engelsk ordkunskap och läsförmåga hos gymnasieelever. Jag söker nu lärare som vill delta med sina klasser. Jag söker åk 1-elever under höstterminen 2019 och åk 3-elever under vårterminen 2020 där de besvarar en bakgrundsenkät och gör två ordförrådstest, ett i svenska och ett i engelska.

Varför är undersökningen relevant?
Syftet med denna delstudie är att skapa kunskap kring hur gymnasieelevers akademiska ordförråd i engelska och svenska ser ut. Ordförråd har visat sig vara en av de viktigaste nycklarna till läsförståelse. Ordförrådet som testas är ord som kan sägas vara bra att kunna när man läser texter som är på en svårare och mer akademisk nivå oavsett ämne.

Er skola har blivit tillfrågad eftersom den har studieförberedande program. En skola kan välja att delta med enskilda klasser upp till hela årskullar, men grundkravet är att samma program erbjuder både elever från åk 1 och åk 3 för att visa på progression.

Hur går det till?
Tiden för enkät och de två ordtesten är cirka 60 minuter inklusive information. Bakgrundsenkäten handlar bland annat om gymnasieprogram, modersmål, föräldrars bakgrund, och språkliga aktiviteter utanför skolan (såsom till exempel datorspelande och läsande). Ordtesten är ett flervalsformat som består av ett antal ord som ska matchas med synonymer/förklaringar. Både enkät och test sker med papper och penna, som jag tillhandahåller när jag kommer ut till klasserna.

För att underlätta logistiken på skolan kan alla elever genomföra materialen och på så sätt ges tillfälle till lärande, erfarenhet och reflektion. Allt deltagande är dock frivilligt, sker på skoltid, ska samtyckas till av eleverna. Eleverna kommer att informeras i förväg, få ställa frågor när jag kommer och få ge sitt samtycke. De kan även ändra sitt samtycke efter genomförande. Undervisande lärare kan välja att vara närvarande eller inte när jag möter klassen.

All information kommer att avidentifieras, det vill säga inget av det eleverna skriver eller kryssar i kommer att publiceras med koppling till just dem eller till skolan. All anonymiserad data sparas enligt gällande principer för datahantering och sekretess.

Återkoppling
Inga resultat på individnivå kan återkopplas, men ett resultat på gruppnivå kan presenteras för läraren/rektorn där man bl.a. får information om ingångs- och utgångsvärden.

Hör av Er på mejl om Ni vill delta med Er/a klass/er och program eller har frågor så återkommer jag med mer information!

Med vänlig hälsning,
Marcus Warnby

E-post: Marcus.warnby@isd.su.se
https://www.su.se/profiles/mawa0203
Institutionen för språkdidaktik, SU
www.isd.su.se

Appendix B

Stockholms universitet

SAMTYCKE TILL DELTAGANDE I STUDIE

Bästa/e gymnasieelev!

Jag heter Marcus Warnby är forskarstudent vid Stockholms universitet. Jag studerar just nu svenskt och engelskt ordförråd hos gymnasieelever. Jag undersöker också om saker utanför skolan kanske påverkar en persons ordförråd. För att svara på detta använder jag en enkät och två ordförrådstest.

Deltagande i studien är frivilligt. Allt deltagande sker på skoltid och alla elever genomför testen och enkäten. Om du inte samtycker kommer materialet att rivas av mig. Om du efteråt kommer på att du inte vill delta i studien är det helt okej att meddela detta och då rivs pappren.

All information kommer att avidentifieras, det vill säga inget av det du skriver eller kryssar i kommer att publiceras med koppling till just dig. Alla svar, utom personnummer, kommer att matas in i dator. Det insamlade materialet, enkät och ordförrådstesten kommer att hanteras enligt rådande arkiverings- och sekretesslag. Ett deltagande innebär att man godkänner att jag (forskaren) får använda informationen och hämta ut data ur utbildningsregistret från SCB för att jämföra med till exempel provbetyg i grundskolan. Delar av materialet kan komma att presenteras i vetenskapliga sammanhang och dokumenteras i en doktorsavhandling.

Ditt deltagande påverkar inte ditt betyg. Din/a lärare får inte återkoppling på just dina test. Däremot kan lärarna få en återkoppling från mig på gruppnivå, det vill säga hur resultaten ser ut för alla elever på skolan.

Tiden för enkät och de två ordtesten är 60 minuter inklusive information. Bakgrundsenkäten handlar om bland annat modersmål, föräldrars utbildning, språkliga aktiviteter utanför skoltid, m.m. Orden testas med ett alternativsvar där ord ska matchas med synonymer/förklaringar. Både enkät och test sker med papper och penna.

En del av orden kommer att vara svåra, men gör så gott du kan!

Om du godkänner till att delta ber jag dig att **fylla i rutorna nedan** genom att kryssa i att du samtycker samt skriva in ditt personnummer.

☐ Jag har fått muntlig och skriftlig informationen om studien och har haft möjlighet att ställa frågor.

☐ Jag samtycker till att delta i studien och att uppgifter om mig behandlas på det sätt som beskrivits

Personnummer: __ __ __ __ __ __ - __ __ __ __

Namn: _____

Mejladress: _____

Lycka till!

Marcus Warnby
Marcus.warnby@isd.su.se

Appendix C

Elevenkät

I det här häftet finns frågor om dig själv och dina vårdnadshavare. En del frågor handlar om fakta medan andra frågor handlar om dina vanor och åsikter.

På en del frågor ska du skriva ett svar, men de flesta frågor har ett antal svarsalternativ. Kryssa i rutan bredvid det svarsalternativ som du väljer. Man kryssar alltså i endast **en** ruta. Här är två exempel:

Exempel 1.
Går du i skolan? ☒ ja ☐ nej

Exempel 2

Hur ofta gör du följande saker?	*Aldrig eller nästan aldrig*	*En eller två gånger i månaden*	*En eller två gånger i veckan*	*/Nästan/ varje dag mindre än 2 tim*	*/Nästan/ varje dag mer än 2 tim*
1) Jag pratar med kompisar	☐	☐	☐	☐	☒
2) Jag tränar idrott	☐	☐	☒	☐	☐
3) Jag åker skateboard	☒	☐	☐	☐	☐

Bakgrundsinformation

Biologiskt kön: ☐ flicka ☐ pojke ☐ vill ej uppge
Upplevt kön: ☐ flicka ☐ pojke ☐ annat ☐ vill ej uppge

Födelseland. I vilket land är du född? ☐ Sverige ☐ Annat: vilket? _____
Om annat, är du adopterad? ☐ ja ☐ nej

Hur länge har du bott i Sverige?
- ☐ 1-5 år
- ☐ 6-10 år
- ☐ mer än 10 år, men inte hela mitt liv
- ☐ Hela mitt liv

Hur många år har du gått i svensk grundskola?
- ☐ 1-4 år
- ☐ 5-8 år
- ☐ hela grundskolan
- ☐ aldrig gått i svensk skola före gymnasiet

Vilket eller vilka är ditt/dina modersmål?
- ☐ Svenska
- ☐ Annat: _____

Om du har annat modersmål än svenska:
a) Har du haft hemspråks/modersmålsundervisning i skolan? ☐ ja ☐ nej
b) Vilken kursplan i svenska får du betyg i? ☐ svenska ☐ svenska som andraspråk

Använder du något annat språk än svenska på fritiden? I så fall, vilket/vilka? Med vem?
Familj: _____

Vänner: _____

Vilket språk tycker du att du är bäst på läsa? _____

Hur många år har du läst engelska i skolan?
☐ 1-3 år ☐ 4-6 år ☐ 7-9 år ☐ 10 år eller mer

Appendix D

I det här häftet finns frågor om dig själv och dina vårdnadshavare. En del frågor handlar om fakta medan andra frågor handlar om dina vanor och åsikter.

På en del frågor ska du skriva ett svar, men de flesta frågor har ett antal svarsalternativ. Kryssa i rutan bredvid det svarsalternativ som du väljer. Man kryssar alltså i endast **en** ruta. Här är två exempel:

Exempel 1.

Går du i skolan? ☒ ja ☐ nej

Exempel 2.

Hur ofta gör du något av följande?	*Aldrig eller Nästan aldrig*	*1-2 gånger varje månad*	*1-2 gånger varje vecka*	*/Nästan/ dagligen mindre än 2 tim*	*/Nästan/ dagligen mer än 2 tim*
1) Jag pratar med kompisar	☐	☐	☐	☐	☒
2) Jag läser böcker	☐	☐	☒	☐	☐
3) Jag spelar datorspel	☒	☐	☐	☐	☐

- Match the definition on the left with the best corresponding word on the right. Write the number
- (1-6) on the line.
- Work as quickly as you can.
- You will find many words difficult, but do your very best!
- If you have no idea about the meaning of a word, do not guess! But if you think you might know the meaning, then you should try to find the answer.

Example:

___6___ part of a house	1 business
	2 clock
___3___ animal with four legs	3 horse
	4 pencil
___4___ something used for writing	5 shoe
	6 wall

Förälder 1/Vårdnadshavare 1.

Vilken är den högsta utbildning som du vet hen har? (kryssa i ett alternativ)

- ☐ Har inte gått i skolan eller inte hela grundskolan
- ☐ 9-årig grundskola
- ☐ Gymnasieutbildning
- ☐ Universitetsstudier upp till 3 år
- ☐ Universitetsstudier mer än 3 år
- ☐ vet inte

Bakgrundsinformation

Biologiskt kön: ☐ flicka ☐ pojke ☐ vill ej uppge

Upplevt kön: ☐ flicka ☐ pojke ☐ annat ☐ vill ej uppge

Födelseland. I vilket land är du född? ☐ Sverige ☐ Annat: vilket? _____

Om annat, är du adopterad? ☐ ja ☐ nej

Hur länge har du bott i Sverige?

☐ 1 - 5 år ☐ 6 - 10 år ☐ mer än 10 år, men inte hela mitt liv ☐ Hela mitt liv

Hur många år har du gått i svensk grundskola?

☐ 1 - 4 år ☐ 5-8 år ☐ hela grundskolan ☐ aldrig gått i svensk skola före gymnasiet

Vilket eller vilka är ditt/dina modersmål? (flera svar är möjliga)

☐ Svenska ☐ Annat: _____

Om annat än svenska: Har du haft hemspråks/modersmålsundervisning i skolan? ☐ ja ☐ nej

Använder du något annat språk än svenska på fritiden? I så fall, vilket/vilka? Med vem?

Familj: _____

Vänner: _____

Vilket språk tycker du att du är bäst på läsa? _____

Hur många år har du läst engelska i skolan? ☐ 1 - 3 år ☐ 4-6 år ☐ 7-9 år ☐ 10 år eller mer

Har du vistats en längre tid i ett engelskspråkigt land? *(språkresa, utbytesår, bosatt…)*

☐ Ja

☐ Nej

→ Om ja, hur länge har du totalt bott i ett engelskspråkigt land?

☐ 2-3 veckor

☐ 1-6 månader

☐ 7-12 månader

☐ mer än ett år

Modernt språk i grundskolan eller gymnasiet?

☐ Ja, det har jag fortfarande

☐ Ja, det har jag haft

☐ Nej, det har jag aldrig haft

→ Om ja, vilket eller vilka?

☐ franska

☐ spanska

☐ tyska

☐ kinesiska/mandarin

☐ annat: _____

Ser du dig som flerspråkig? ☐ Ja ☐ Nej

Tänker du att du ska studera på högskola/universitet direkt efter gymnasiet? ☐ Ja ☐ Nej ☐ Vet inte

Du ska nu kryssa dina tidigare betyg i Engelska. Kryssa i det du minns eller tror att du hade.

Grundskolan, Årskurs 9?	☐ A	☐ B	☐ C	☐ D	☐ E	☐ F	☐ Inget betyg
Gymnasiet, Engelska 5?	☐ A	☐ B	☐ C	☐ D	☐ E	☐ F	☐ Inget betyg
Gymnasiet, Engelska 6?	☐ A	☐ B	☐ C	☐ D	☐ E	☐ F	☐ Inget betyg

Läser du kursen Engelska 7 nu i årskurs 3? ☐ Ja ☐ Nej

Hur skulle du bedöma din läsförmåga i **engelska**?

- □ Svag
- □ Under genomsnittet
- □ Okej
- □ Bra
- □ Utmärkt

Hur snabbt tycker du att du läser **engelsk** text som hör till skolarbetet?

- □ Väldigt långsamt
- □ Ganska långsamt
- □ Genomsnittligt
- □ Ganska snabbt
- □ Väldigt snabbt

Hur tycker du att det är att förstå **engelsk** text som hör till skolarbetet?

- □ Mycket är svårt att förstå
- □ Ganska mycket är svårt att förstå
- □ En del saker är lätta, andra saker är svåra att förstå
- □ De flesta saker är lätta
- □ Allt är lätt att förstå

* * *

Du ska nu ge information om dina föräldrar

Förälder/Vårdnadshavare 1.
Biologiskt kön. □ kvinna □ man □ vill ej uppge

Vilken är den högsta utbildning som du vet hen har? (kryssa i ett alternativ)

- □ Har inte gått i skolan eller inte hela grundskolan
- □ 9-årig grundskola
- □ Gymnasieutbildning
- □ Universitetsstudier upp till 3 år
- □ Universitetsstudier mer än 3 år
- □ vet inte

Vilket eller vilka är hens modersmål? □ Svenska *och/eller* □ Annat: _____

I vilket land är hen född? □ Sverige □ Annat: vilket? _____

Förälder/Vårdnadshavare 2.
Biologiskt kön. □ kvinna □ man □ vill ej uppge

Vilken är den högsta utbildning som du vet hen har? (kryssa i ett alternativ)

- □ Har inte gått i skolan eller inte hela grundskolan
- □ 9-årig grundskola
- □ Gymnasieutbildning
- □ Universitetsstudier upp till 3 år
- □ Universitetsstudier mer än 3 år
- □ vet inte

Vilket eller vilka är hens modersmål? □ Svenska *och/eller* □ Annat: _____

I vilket land är hen född? □ Sverige □ Annat: vilket? _____

Fritt skrivande
Du går nu sista året på gymnasiet i ett högskoleförberedande program. Känner du att du börjar bli förberedd på att läsa och studera på universitet?

Brukar ni läsa engelska texter i andra ämnen än engelska? Skriv fritt hur du tycker det är att läsa engelska texter i skolan? Brukar ni använda både bild, film och text för att jobba med ett ämne? Hur brukar du göra när det är svårt att förstå engelskan i skoltexterna? Stannar du ofta upp och inte förstår ord? Är det några ämnen där texterna har varit svårare än andra?

Skriv om du tycker att du att lär dig språk när du pratar med andra eller läser eller hör språk?
Tänker du på att du ska lära dig engelska när du ser på film, läser tidningar, spelar datorspel, osv.? Är det skillnad på dina vanor att spela datorspel på engelska, se på engelsk film/TV, och liknande nu när du snart gått tre år på gymnasiet mot när du gick i högstadiet? Hur i så fall?

Appendix E

Språkliga aktiviteter utanför skoltid

På en skala från 1 – 5 ska du nu bedöma ungefär hur mycket du läser, ser och hör engelska i vardagen. Försök att svara hur du brukar göra i genomsnitt.

Hur ofta gör du följande saker? **1. Jag tittar på …**	*Aldrig eller nästan aldrig*	*En eller två gånger i månaden*	*En eller två gånger i veckan*	*/Nästan/ varje dag mindre än 2 tim*	*/Nästan/ varje dag mer än 2 tim*
a) svensktextad film/serier där man talar engelska	☐	☐	☐	☐	☐
b) otextad film/serier där man talar engelska	☐	☐	☐	☐	☐
c) engelsktextad film/serier där man talar engelska	☐	☐	☐	☐	☐
d) engelsktextad film/serier där man talar annat språk	☐	☐	☐	☐	☐
2. Jag tittar på engelska faktaprogram (t.ex. dokumentärer, nyheter, reportage) som är…					
a) svensktextade	☐	☐	☐	☐	☐
b) otextade	☐	☐	☐	☐	☐
c) engelsktextade	☐	☐	☐	☐	☐
3. Jag tittar på andra engelsktalade program (t.ex. reality/dokusåpor, youtubeklipp, v-loggar) som är…					
a) svensktextade	☐	☐	☐	☐	☐
b) otextade	☐	☐	☐	☐	☐
c) engelsktextade	☐	☐	☐	☐	☐
4. Jag läser skönlitterära texter på engelska (t.ex. romaner, noveller, deckare. Digitalt eller i pappersform)	☐	☐	☐	☐	☐
5. Jag läser faktatexter på engelska (t.ex. om teknik, historia, mode, sport. Digitalt eller i pappersform)	☐	☐	☐	☐	☐
6. Jag läser nyhetstexter på engelska (t.ex. artiklar, dagstidningar digitalt eller i pappersform)	☐	☐	☐	☐	☐
7. Jag lyssnar på musik där de sjunger på engelska	☐	☐	☐	☐	☐
8. Jag lyssnar på radio/podd där de talar engelska	☐	☐	☐	☐	☐
9. Jag spelar digitala spel där språket är engelska	☐	☐	☐	☐	☐

Om du på fråga 9 svarat att du spelar ganska mycket digitala spel:

a) Hur mycket spelar du spel där du talar/chattar med andra?
☐aldrig ☐ 1–5 tim/v ☐ 6–10 tim/v ☐11–15 tim/v ☐16 tim/v eller mer

b) När du spelar onlinespel där spelet är på engelska, talar/chattar du då:
☐ Mest på svenska ☐mest på engelska ☐ungefär lika mycket på svenska som på engelska ☐Annat: _____

Appendix F

31 How often do you come into contact with [target language] through media in the following ways?

(Please select one answer from each row)

	Never	*A few times a year*	*About once every month*	*A few times a month*	*A few times a week*
1) How often do you listen to songs in [target language]?	○ 0	○ 1	○ 2	○ 3	○ 4
2) How often do you watch movies spoken in [target language] without subtitles?	○ 0	○ 1	○ 2	○ 3	○ 4
3) How often do you watch movies spoken in [target language] with subtitles?	○ 0	○ 1	○ 2	○ 3	○ 4
4) How often do you watch television programmes (not movies) spoken in [target language] without subtitles?	○ 0	○ 1	○ 2	○ 3	○ 4
5) How often do you watch television programmes (not movies) spoken in [target language] with subtitles?	○ 0	○ 1	○ 2	○ 3	○ 4
6) How often do you play computer games spoken in [target language]?	○ 0	○ 1	○ 2	○ 3	○ 4
7) How often do you read books written in [target language]?	○ 0	○ 1	○ 2	○ 3	○ 4
8) How often do you read a magazine or a comic written in [target language]?	○ 0	○ 1	○ 2	○ 3	○ 4
9) How often do you visit websites written in [target language]?	○ 0	○ 1	○ 2	○ 3	○ 4

Appendix G

ST176

How often are you involved in the following reading activities?

(Please select one response in each row. If you don't know what the activity is, please select "I don't know what it is.")

	I don't know what it is	Never or almost never	Several times a month	Several times a week	Several times a day
Reading emails	○	○	○	○	○
Chat online (e.g., WhatsApp®, Messenger®)	○	○	○	○	○
Reading online news	○	○	○	○	○
Searching information online to learn about a particular topic	○	○	○	○	○
Taking part in online group discussions or forums	○	○	○	○	○
Searching for practical information online (e.g., schedules, events, tips, recipes)	○	○	○	○	○

Index