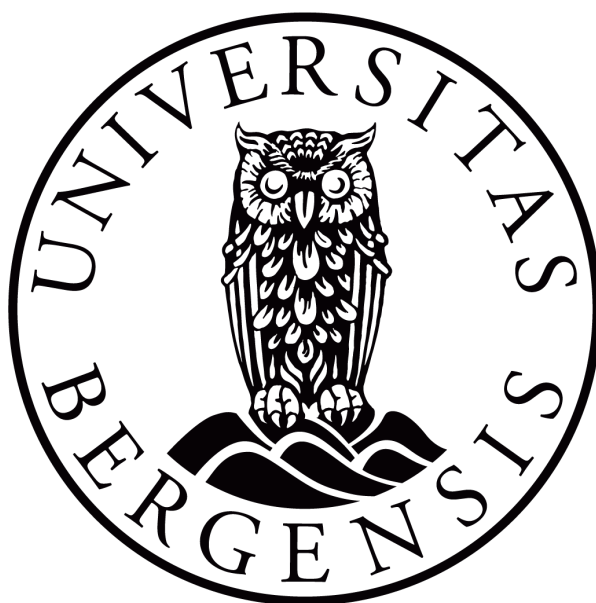


# **Academic Vocabulary Production**

A Corpus Study of Written Texts by Norwegian L2 Learners of English in  
Upper Secondary Education

**Filip Thorvaldsen**



Master thesis

Department of Foreign Languages

University of Bergen

May 2024

## Abstract in Norwegian

Akademiske ord utgjør en viktig del av et akademisk språk. Kunnskap til slike ord har blitt knyttet til akademisk oppnåelse. En tendens har vært å fokusere på elever og studenter sin kunnskap knyttet til reseptive ferdigheter, hovedsakelig lesing. Produktiv bruk av akademiske ord i skriftlige tekster anses som en vanskeligere ferdighet. Studier har belyst at et akademisk ordforråd kan påvirke vurderingen av skriftlig kvalitet i positiv retning, og at det er en viktig ferdighet i høyere utdanning. På bakgrunn av den økte innflytelsen av engelsk i høyere utdanningsinstitusjoner, er det viktig at også norske elever lærer seg engelske akademiske ord.

Akademiske ord er definert i denne studien som generelle og er operasjonalisert som ordene på the Academic Vocabulary List (AVL). Målet med denne studien var å undersøke bruken av disse ordene i skriftlige tekster blant norske elever. For å undersøke dette laget jeg et korpus bestående av 28 engelsk tentamenstekster skrevet av norske elever, første året på videregående, i et studiespesialiserende program.

Resultatene viser at nesten 8% av tekstene består av akademiske ord, i gjennomsnitt. Dette er omtrent tilsvarende omfanget av akademiske ord som finnes i avistekster. Majoriteten av de akademiske ordene stammer fra kun en liten del av AVL ordlisten. I tillegg er de fleste ordene også å finne i forberedelsesmateriale og oppgaveheftene til engelsk tentamen. I alt produserte elevene akademiske ord i varierende grad. Videre fant studien at kun et fåtall akademiske ord ble brukt feil semantisk. Blant disse tyder funnene imidlertid på at formlighet mellom ordet som ble brukt feil og det riktige ordet, kan ha vært en avgjørende faktor. Til sist fant studien at sammenhengen mellom skriftlig akademisk ordbruk og oppnådd karakter ikke var betydelig stor. Likevel viser den deskriptive statistikken at elevene med høyest karakter, i gjennomsnitt, brukte en noe større andel akademiske ord.

Funnene indikerer at et økt fokus på akademiske ord i engelskundervisningen kan være hensiktsmessig, på bakgrunn av variasjonen blant elevene, og særlig for ord med stor formlighet. I tillegg tyder funnene på at lærere kan benytte seg av tekster elever må lese i skriveprosessen, som en ressurs for å hjelpe elevene med å produsere akademiske ord.

## **Acknowledgements**

I would like to thank my supervisor, Kimberly Skjelde, for valuable guidance. I would also like to thank the learners who participated in this study, and the teachers who agreed to help in this process.

## Table of Contents

Abstract in Norwegian .....	ii
Acknowledgements .....	iii
List of Tables and Figures .....	vi
List of Abbreviations.....	vii
1.0 Introduction .....	1
1.1 Academic Vocabulary Knowledge in SLA .....	1
1.2 The Norwegian Educational Context .....	3
1.3 Written Production .....	4
1.4 Corpus-Based Study .....	6
1.5 Aim and Scope .....	6
1.6 Research Questions .....	7
1.7 Outline of the Thesis .....	7
2.0 Theoretical Background .....	8
2.1 Usage-Based Theory .....	8
2.1.1 Vocabulary Acquisition in Usage-Based Approaches .....	10
2.1.2 Frequency .....	10
2.1.3 Salience and Noticing.....	12
2.1.4 The Implicit and Explicit Dichotomy.....	13
2.2 Corpus Linguistics.....	16
2.2.1 General and Specialized Corpora .....	17
2.2.2 Word Count Units .....	19
2.3 Vocabulary Knowledge.....	20
2.3.1 Receptive and Productive Knowledge .....	20
2.3.2 Vocabulary and Writing .....	21
2.4 Academic vocabulary .....	25
2.5 Word lists .....	27
2.5.1 The Academic Word List (AWL) .....	28
2.5.2 The Academic Vocabulary List (AVL).....	29
2.6 Previous Research on Productive General Academic Vocabulary .....	35
2.6.1 General Academic Vocabulary Production in Written Texts.....	36
2.6.2 Factors Impacting General Academic Vocabulary Use: Source Texts, Topic, and Task .....	39
2.6.3 General Academic Vocabulary Use and Assessment of Writing.....	41
3.0 Methods and Procedures .....	43
3.1 Research Design .....	43
3.2 Participants .....	44

3.3 Material .....	46
3.3.1 Corpus Construction.....	46
3.4 Data Analysis .....	51
3.4.1 Profiling the Corpus for General Academic Vocabulary .....	51
3.4.2 Examining the Impact of Assignment Briefs and Preparation Material .....	56
3.4.3 Evaluating Correct Use .....	58
3.4.4 Investigating the Relationship with Achieved Grades .....	59
3.5 Validity and Reliability .....	60
3.6 Ethics.....	63
4.0 Results .....	64
4.1 General Academic Vocabulary Production in the Corpus .....	64
4.2 Impact of Assignment Briefs and Preparation Material.....	67
4.3 Correct Use of General Academic Vocabulary.....	69
4.4 Relationship with Achieved Grades.....	71
5.0 Discussion of Results .....	75
5.1 General Academic Vocabulary Production in the Corpus .....	75
5.1.1 AVL Vocabulary Coverage.....	75
5.1.2 Variation in AVL Use .....	77
5.1.3 AVL Vocabulary from the Top 500 Words on the List .....	78
5.1.4 The non-AVL vocabulary .....	79
5.1.5 Potential for Development .....	80
5.2 Impact of Assignment Briefs and Preparation Material.....	82
5.3 Correct Use of General Academic Vocabulary.....	84
5.4 Relationship with Achieved Grades.....	88
6.0 Conclusion.....	90
6.1 Key findings .....	90
6.2 Implications .....	91
6.2.1 Increased Focus on General Academic Vocabulary .....	91
6.2.2 Vocabulary Learning Potential in Reading Texts when Writing .....	93
6.3 Limitations .....	94
6.3.1 Methodological Limitations .....	94
6.3.2 Limitations with Single Words .....	95
6.4 Recommendations for Further Research .....	96
References .....	98
Appendices.....	112

## List of Tables and Figures

Table 1. Overview of Distinctions in Corpus Linguistics (Adapted from McEnery and Hardie, 2012, p. 3) .....	17
Table 2. Aspects of Productive Vocabulary Knowledge Needed for Writing (Adapted from Nation, 2022, p. 54, with Inspiration from Coxhead, 2012, p. 138) .....	22
Table 3. Top 20 AWL Word Families, AVL Word Families, and AVL Lemmas .....	31
Table 4. Composition of Tokens in the Corpus .....	48
Table 5. Overview of Task Differences, with Recommended Word Limit in Parenthesis .....	49
Table 6. Overall Corpus Composition and AVL Coverage .....	64
Table 7. Distribution of AVL Tokens in the Corpus in Percentages .....	65
Table 8. Distribution of AVL Tokens in the Corpus in Raw Numbers .....	66
Table 9. AVL Coverage for Each Task Making up the Mock Exams .....	66
Table 10. AVL Tokens Achieved by AVL Types from the Assignment Briefs .....	67
Table 11. AVL Tokens Achieved by AVL Types from the Assignment Briefs and Preparation Material Combined .....	68
Table 12. Correct Use of AVL Words Produced in the Corpus in Percentages .....	69
Table 13. AVL Use in Percentages in the Two Groups Based on Grades .....	71
Figure 1. Example of Profiling a Text in AntWordProfiler .....	53
Figure 2. AVL Token Coverage of the Corpus .....	65
Figure 3. Scatter Plot of AVL Use by Grade .....	72
Figure 4. Boxplot of AVL Use by Grade .....	73

## List of Abbreviations

AVL - The Academic Vocabulary List (Gardner & Davies, 2014)

AWL - The Academic Word List (Coxhead, 2000)

BAWE - British Academic Written English

BNC - British National Corpus

COCA - Corpus of Contemporary American English

CLIL - Content and Language Integrated Learning

EAP - English for Academic Purposes

EFL - English as a Foreign Language

EMI - English Medium Instruction

ESL – English as a Second Language

FL - Foreign Language

GSL - General Service List (West, 1953)

LCR – Learner Corpus Research

L1 - First Language

L2 - Second Language

POS – Part of Speech

RQ(s) – Research questions(s)

SLA - Second Language Acquisition

UWL – The University Word List (Xue & Nation, 1984)

## 1.0 Introduction

“People and language create each other, grow from each other, and act and change under the influence of the other.”

Nick C. Ellis (2015, p. 49)

### 1.1 Academic Vocabulary Knowledge in SLA

Vocabulary, or lexis, is an important aspect of Second Language Acquisition (SLA). Without the ability to understand and produce words, any participation in a second language (L2) discourse, will be challenging. However, different types of vocabulary are important for various purposes. The first 2,000-3,000 most frequent words in English are important for general use and communication. After having acquired this, more specialized vocabulary deserves attention (Nation, 2022). For learners aiming toward a specific trade or vocation, technical vocabulary is useful to learn, which has a close relationship to the content of certain disciplines. However, for learners aiming toward tertiary education, another type of specialized vocabulary is important, irrespective of discipline. This is called general academic vocabulary, which is the topic of this thesis. This vocabulary is defined as words used more frequently in academic language than in non-academic language, and is cross-disciplinary (Nagy & Townsend, 2012). In fact, studies have shown that this lexis makes up between 10% (Coxhead, 2000) to 14% (Gardner & Davies, 2014) of academic texts and considerably less in fiction or newspapers, which is one of the reasons why this vocabulary is important. It is also important to acquire this lexis as it entails learning the tools necessary to think in academic terms (Nagy & Townsend, 2012). The current study focuses on Norwegian upper secondary learners' production of general academic vocabulary in English.

Academic words are often defined by their inclusion on a word list (Schmitt & Schmitt, 2020). Several lists have been created, such as the widely used Academic Word List (AWL) (Coxhead, 2000), and the more recent Academic Vocabulary List (AVL) (Gardner & Davies, 2014). Based on principles and methods in corpus linguistics, such lists have been constructed from corpora containing authentic language use (Dang, 2020). Corpora are widely used in both SLA (Römer, 2013) and vocabulary research (Schmitt, 2010). In the creation of general academic vocabulary lists, corpora of texts representative of academic discourse have been utilized (Coxhead, 2000; Gardner & Davies, 2014). Words occurring more frequently in such corpora, compared to corpora comprised of general texts, are usually the ones included in the list (Schmitt & Schmitt, 2020).



Much of academic vocabulary is Graeco-Latin and can be challenging to learn because they do not appear frequently in the language, they are often abstract, and the meaning of the words might not be revealed by their forms (Nation, 2022). Instead, this vocabulary largely contributes to the rhetoric (Paquot, 2010), precision and tone (Schmitt & Schmitt, 2020), which signals “the formal academic nature of a text” (Nation, 2022, p. 309). Because this lexis is not related to the topic or content of a text, but rather supports it, it has low salience (Coxhead, 2000). This can make it difficult to notice. Indeed, learners often struggle more with general academic vocabulary than with technical vocabulary (Nation, 2022). Another problem is that L2 learners of English might rarely encounter general academic vocabulary outside of school, in their spare time (Olsson & Sylvén, 2015). As a result of the limited exposure to this vocabulary in everyday communication, Corson (1995) introduced the term “lexical bar.” This term is defined as representing:

a gulf between the everyday meaning systems and the high status meaning systems created by the introduction of an academic culture of literacy. This is a barrier that everyone has to cross at some stage in their lives, if they are to become ‘successful candidates’ in conventional forms of education. (Corson, 1995, pp. 180-181)

Furthermore, in a subsequent article, Corson (1997) argues that this lack of experience with academic vocabulary might be particularly true for learners from certain socio-cultural backgrounds. If learners rarely meet this lexis in everyday life, it is difficult to use it in the educational system (Corson, 1997).

Therefore, teachers have a responsibility to lower this lexical bar, especially considering that knowledge of general academic vocabulary has been linked to academic achievement and success (Corson, 1997; Coxhead, 2000; Gardner & Davies, 2014; Nagy & Townsend, 2012). Transitioning from high school to university entails using language in novel ways, and both natives as well as non-natives are required to undergo this adjustment (Biber, 2006). Indeed, academic language is not the native or first language for any learner of English (Mauranen et al., 2010; Tang, 2012). However, for L2 learners of English, academic language might be more difficult to cope with, than for native speakers (Snow & Uccelli, 2009). A crucial element of what constitutes academic language is academic vocabulary (Nagy & Townsend, 2012).

Furthermore, the role of English in tertiary education and academia has become more prominent. In fact, English might be referred to as “the global lingua franca of academia” (Mauranen et al., 2010, p. 183). One consequence of this development is that English is increasingly being used in the higher education system in countries where another language is

used at lower levels, such as primary and secondary school (Henriksen et al., 2019). In other words, the language of tertiary education is academic in nature, and English is gradually turning into the lingua franca.

## **1.2 The Norwegian Educational Context**

This trend has also been documented in Norway, where the role of English is increasing within higher education (Ljosland, 2007). In the English subject curriculum in Norway (LK20), the role of English as a global language is emphasized in the relevance and central values of the subject, where it states that English “shall prepare the pupils for an education and societal and working life that requires English-language competence in reading, writing and oral communication” (Ministry of Education and Research, 2019, p. 2). Considering the amount of English teaching Norwegian learners have received before entering tertiary education, one might expect them to be prepared for this.

English is a compulsory subject from first until tenth grade in the Norwegian educational system. After having finished ten obligatory years of education, learners will have had 588 hours of English instruction (Norwegian Directorate for Education and Training, 2023). Although upper secondary education is not mandatory in Norway, for those who choose it, English is compulsory in the first year. Upon completion of this year, learners will have had an additional 140 hours of English (Norwegian Directorate for Education and Training, 2023). Norwegians only need to complete the first year of English in upper secondary school to study at university and not pass an English proficiency test like international students (Norwegian Universities and Colleges Admission Service, 2024). Thus, when this first year is completed, these learners will not receive any more English teaching before they enter tertiary education, unless they choose English as an elective for the next two years. Skarpaas (2011) found that the majority (55%) of Norwegian learners in their final year of upper secondary education, in a general studies programme, had not chosen English as an elective.

Therefore, one could say that “Norwegian institutions of higher education take for granted that English as a foreign language (EFL) instruction in upper-secondary schools effectively prepares students for the use of English in higher education” (Hellekjær, 2009, p. 199). Indeed, Norwegian learners have been shown to be proficient in English compared to other countries (Bonnet, 2004). In the most recent publication of the EF English Proficiency Index (EPI) (Education First, 2023), Norway ranked fifth on a list of 113 countries and

regions. In fact, since the first EF EPI was published in 2011, Norway has never been ranked below fifth place. Although these tests only measure people who are 18 years and older, they indicate a high English proficiency among Norwegians. Nonetheless, the level of English proficiency required in tertiary education may be more difficult to acquire. In fact, Hellekjær (2005, 2008, 2009) found that students in Norway were not proficient enough to read texts and textbooks written in English in tertiary education. Similar results were found by Busby (2015), whose results suggest that Norwegian students are proficient in general English, but “appear to have difficulties with the language of academia” (p. 79). This indicates that general studies may not sufficiently prepare the learners to participate in the academic discourse in higher education.

In studies of learner language, there has been a tendency to study learners in a foreign rather than second language environment (Paquot & Plonsky, 2017). Rindal and Brevik (2019) point out that both labeling English as a second language (ESL) and as a foreign language (EFL) is common in Norway. However, the authors underline that both labels can be problematic, and suggest the use of a more generic term “L2 English ... [which] refers to English as *a second or later language*” (Rindal & Brevik, 2019, p. 435, emphasis in original). In the present study, I use the same label. Furthermore, the majority of previous research into learner language has focused on university students, while learners in secondary education have received less attention (Paquot & Plonsky, 2017). This study aims to fill this gap by investigating L2 learners in upper secondary education.

In Norway, one of the competence aims after the first year in upper secondary school, for general studies, states that the L2 learner is expected to “listen to, understand, and use academic language in working on one’s own oral and written texts” (Ministry of Education and Research, 2019, p. 12). In this competence aim, both receptive and productive abilities connected to academic language, are emphasized.

### **1.3 Written Production**

In previous research on general academic vocabulary much of the focus has been placed on receptive knowledge (Durrant, 2016; Schmitt, 2010). This might be because L2 learners of English likely will encounter situations requiring them to read or listen to academic discourse in English, at university. However, university students do not only need receptive knowledge. They also need productive knowledge in many cases, such as when writing essays, reports, MA theses, and so on (Paquot, 2010). In addition, most assessment in tertiary education is

achieved through writing (Coxhead, 2012). Furthermore, the ability to produce texts according to the conventions of academic writing, is crucial. In fact, vocabulary choice can indicate the extent to which a “writer has adopted the conventions of the relevant discourse community” (Nation, 2022, p. 226). Moreover, the ability to write in a style that is appropriate according to the academic context when studying, can impact students’ academic success (Coxhead & Byrd, 2007). In other words, learning to produce academic vocabulary can be conducive for academic success (Nation, 2022).

Thus, while a receptive focus on general academic vocabulary is important, a productive focus is also necessary. Having receptive knowledge of a word is not the same as being able to produce it (Laufer, 2005). Therefore, measuring or monitoring the size of learners’ productive academic vocabulary can be necessary to gain insight into how prepared they are for the various productive abilities required when studying, such as writing academic papers (Lin & Morrison, 2010). Nonetheless, the ability to produce a word on a test that elicits it does not mean a learner will be able to or choose to use it, in their writing (Laufer, 2005). Measuring vocabulary in free production has been identified as one of the prominent gaps in vocabulary studies (Schmitt, 2010). Examining the extent to which general academic vocabulary is produced in written texts is one of the main objectives for this study.

One of the basic skills in the English subject curriculum (LK20) for Norwegian learners, is the ability to write. Regarding writing proficiency, the curriculum states that “[t]he development of writing proficiency in English progresses from learning single words and phrases to creating different types of coherent texts that present viewpoints and knowledge” (Ministry of Education and Research, 2019, p. 4). In this passage, vocabulary knowledge is highlighted as part of what constitutes writing proficiency. Even though writing consists of more than producing individual words only, the vocabulary choices of learners greatly impact their written texts. Research suggests that learners tend to overuse core, general words (e.g. Hasselgren, 1994; Henriksen & Danelund, 2015). In contrast to general vocabulary, academic words help make a text more concise and precise, as well as to organize the text and build cohesion (Schmitt & Schmitt, 2020). Thus, if learners overuse general words in their written production, the formal nature and tone of academic texts might be compromised (Nation, 2022; Schmitt & Schmitt, 2020).

## 1.4 Corpus-Based Study

To investigate the written production of general academic vocabulary, a small, specialized corpus of written English mock exam texts was compiled in the present study. These texts can be considered high stakes, as they were graded. In the present study, the grades were collected as well, to examine a potential link with general academic vocabulary production. Research has shown that vocabulary knowledge influences writing performance (Milton et al., 2010; Stæhr, 2008) and the assessment of writing quality (Nation, 2022). For general academic vocabulary in particular, a moderate association has been found between receptive knowledge of this lexis and English course grades in Norwegian upper secondary school (Skjelde & Coxhead, 2020). Moreover, Olsson's (2015) study indicates that the holistic assessment of essays produced by upper secondary learners is influenced by the use of general academic vocabulary. Research investigating this relationship is missing in the Norwegian context.

Access to achieved grades was one benefit of creating my own corpus, instead of utilizing existing corpora of learner language. The fact that the mock exam format has changed in recent years to resemble the exam, which consists of four different tasks, was another reason to create my own corpus. Finally, access to the assignment briefs containing the tasks for the mock exam, as well as the preparation material provided to the students in advance, allowed for a further investigation into a potential link between general academic vocabulary in these texts and in the written mock exams. Previous research has suggested that learners rely on vocabulary from source texts and assignment briefs in the writing process (Leki & Carson, 1997; Milton, 2001; Plakans & Gebriel, 2012). Similar results have been found for general academic vocabulary (Therova, 2021, 2022). It is interesting to examine the extent to which previous findings are corroborated in the present study, in a Norwegian context.

## 1.5 Aim and Scope

The vocabulary I examine is general academic vocabulary and productive knowledge of this lexis. In the present study, this vocabulary is operationalized as the words on the AVL created by Gardner and Davies (2014). Furthermore, corpus linguistics is used to conduct the research investigation into learners' production of general academic vocabulary in written texts. I aim to explore the use of general academic vocabulary in written English texts among Norwegian L2 learners of English in upper secondary education in general studies. First, to investigate the extent to which general academic vocabulary is used. Second, to examine the impact of

the assignment briefs and preparation material on this production. Third, to evaluate if the general academic vocabulary is used correctly. And finally, to investigate the relationship between the extent of general academic vocabulary production and achieved grades on the written English mock exam.

## **1.6 Research Questions**

To guide the present study, I have developed the following research questions (RQs):

1. To what extent do Norwegian L2 learners of English in upper secondary general studies produce general academic vocabulary in written mock exam texts?
  - a. To what extent is this vocabulary production influenced by the words found in the assignment briefs and preparation material for the mock exam?
2. To what extent do these learners use general academic vocabulary correctly?
3. Is there an association between the extent of general academic vocabulary production in the written texts and achieved grades for these learners?

## **1.7 Outline of the Thesis**

The first chapter has introduced the topic of this thesis, its aim and scope, and the RQs guiding the present study. Further, chapter 2 presents the theoretical background, describing usage-based theory, corpus linguistics, vocabulary knowledge, academic vocabulary, and word lists. This chapter also presents empirical research relevant to the topic of productive academic vocabulary in written texts. In chapter 3, the methods and analytical procedures in the current study are described. This includes how participants were recruited, what data were used, how the corpus was created, and what types of analyses were conducted. Further, in chapter 4, the results are presented before they are discussed in chapter 5. Finally, chapter 6 concludes with the key findings, implications, limitations, and recommendations for future research.

## 2.0 Theoretical Background

### 2.1 Usage-Based Theory

SLA is a subfield of applied linguistics (Ortega, 2011). Applied linguistics as a field can be defined as “using what we know about (a) language, (b) how it is learned and (c) how it is used, in order to achieve some purpose or solve some problem in the real world” (Celce-Murcia & Schmitt, 2020, p. 1). The purpose or problem-solving is directed towards second language education, in this study, hence locating it in the subfield of SLA. This can be defined as the field that investigates the acquisition of languages after early childhood (Ortega, 2009). It is not necessarily limited to the study of acquiring a second language only, but any additional language learned later in life.

An important part of SLA is the acquisition of vocabulary. Richards (1976) described vocabulary learning and teaching as a neglected issue in both theoretical and applied linguistics. This neglect was also present in SLA research, according to Meara (1980). He noted that:

This neglect is all the more striking in that learners themselves readily admit that they experience considerable difficulty with vocabulary, and once they have got over the initial stages of acquiring their second language, most learners identify the acquisition of vocabulary as their greatest single source of problems. (p. 221)

The interest in vocabulary acquisition in SLA has been growing since that time (R. Ellis, 2008; Nation, 2011). The importance of vocabulary knowledge cannot be understated, as it is “fundamental to all language use” (Schmitt et al., 2015, p. 213). Compared to grammatical competence, vocabulary knowledge is often valued as more essential to communication (Qian & Lin, 2020). The acquisition of vocabulary is a protracted process, and vocabulary knowledge cumulates over time (R. Ellis, 2008). It may be regarded as an incremental process (Barclay & Schmitt, 2019).

Various theories have been proposed to explain how second languages are acquired and developed. In fact, SLA has been described as a diverse and interdisciplinary field, intersecting with associated fields such as psychology, psycholinguistics, educational science, cognitive linguistics, applied linguistics and corpus linguistics, the latter as both method and theory (Gries, 2008). Along similar lines, Schmitt (2010) discusses prominent gaps in vocabulary studies. One of these gaps is that there is no overall theory of vocabulary acquisition, something he refers to as “the Holy Grail of vocabulary studies” (Schmitt, 2010,

p. 36). The solution to comprehend vocabulary acquisition, Schmitt (2010) maintains, is to combine the strengths of different methodologies and paradigms.

In this study, I take a usage-based approach to second language learning and development. Such approaches have increased in influence over the past two decades, in applied linguistics (Römer, 2023). The term was coined by Ronald W. Langacker in 1987. Usage-based theories can be found in both child language acquisition (e.g. Tomasello, 2003), and in SLA. Similar to the interdisciplinary nature of SLA (Gries, 2008) and vocabulary acquisition (Schmitt, 2010), usage-based linguistics builds on empirical findings from cognitive linguistics, corpus linguistics, psycholinguistics, as well as psychology (N. Ellis, 2017). In fact, Ortega (2015) highlights this interdisciplinarity as one of the reasons for why usage-based perspectives have increased in SLA research. Despite the lack of a single unified usage-based model to language learning, the variety of models are unified by the importance which is placed on actual usage (Tyler & Ortega, 2018a). These theories have in common that they view language learning as a process that is dynamic. In this process, usage is what causes language to be emerged (Verspoor & Schmitt, 2013).

Thus, as a second language learning theory, usage-based theory falls under the umbrella term “emergentism” (Mitchell et al., 2019). Theories falling under this term views SLA as bottom-up. This means that:

learners use general learning mechanisms in order to extract structure and patterns from the language input they are exposed to. From this perspective, formal aspects of language “emerge” from language experience, rather than either being innate or being represented as rigid abstract structures. (Mitchell et al., 2019, p. 129)

There are now multiple approaches and theories within SLA that fall under this label. According to N. Ellis and Wulff (2020), this group is united in their view that learners use general cognitive abilities to learn language from exposure to communicative experience, that is, language use. Put differently, using the language in a meaningful way is how it is learned, which is why it is labelled usage-based (Langacker, 2009). A more inclusive term for usage, however, is experience. That is because usage includes both exposure to input, but also practice in output. A result of this is that learners will inevitably acquire their L2 in distinct ways, due to differences in experience (Verspoor & Smiskova, 2012).

Both theories in cognitive linguistics and psycholinguistics are commonly referred to as usage-based perspectives to language acquisition (N. Ellis, 2002). In the current study, I draw on insights from both of these theoretical fields to inform the discussions on noticing and awareness, two concepts related to salience, as well as learners’ sensitivity to frequency



in the input. These issues and mechanisms are intricately connected to the explicit and implicit learning and knowledge dichotomies, which ultimately have implications for the relationship between explicit and implicit instruction.

### **2.1.1 Vocabulary Acquisition in Usage-Based Approaches**

While usage-based SLA approaches can be applied to the acquisition of a variety of linguistic units, this study focuses on general academic vocabulary. We may therefore begin by highlighting the notion of constructions, which is how words are defined in usage-based approaches. The term is defined by N. Ellis (2012) as “form-meaning mappings, conventionalized in the speech community, and entrenched as language knowledge in the learner’s mind” (p. 197). Other terms for entrenchment are automatization and habit-formation, and essentially refer to how easy it is to retrieve or use a word (Verspoor & Smiskova, 2012). When a learner pairs a form and a meaning, he or she associates a given form, such as a lexical item, with a particular meaning, and the result is considered a construction (N. Ellis & Wulff, 2020). In a usage-based approach to SLA, the acquisition of language can be regarded as learning constructions (Römer, 2023). Another term often used synonymously is conventional units. Referring to Langacker’s (2000) work, Verspoor and Schmitt (2013) highlight that, in a usage-based approach, expressions that have been conventionalized (i.e. conventional units) are what comprise the lexicon.

These constructions or conventional units may range from morphological, to lexical, to syntactic units (N. Ellis, 2002) such as formulaic sequences (Verspoor & Schmitt, 2013). According to Langacker (2008), these units are abstracted from usage events, that is, language use in context. Therefore, vocabulary, academic or otherwise, is considered to be acquired through input, but also through using the language in communication (N. Ellis, 2018). Using vocabulary in a meaningful way, such as through writing, is one way for these constructions to be strengthened (Verspoor & Smiskova, 2012). Another way in which they are strengthened is by frequently encountering them in the input (N. Ellis & Wulff, 2020).

### **2.1.2 Frequency**

Usage-based approaches often view word frequency as an important component in lexical development (Crossley et al., 2016). N. Ellis (2013) draws on evidence from psycholinguistic research in support for a usage-based approach to language learning. He underlines the importance of frequency and emphasizes that “[f]requency is a key determinant of acquisition

because “rules” of language, at all levels of analysis [...] are structural regularities which emerge from learners’ lifetime unconscious analysis of the distributional characteristics of the language input” (N. Ellis, 2013, p. 261). The importance of frequency in usage-based approaches is further emphasized by Kemmer and Barlow (2000) who underline that higher frequency impacts entrenchment. Thus, the more frequent a learner is exposed to a unit, such as a lexical item, the better it is learned. As a consequence it is more easily understood and more readily produced. In usage-based models, frequency is often operationalized in two different ways, as token and type frequency. The former refers to “how often a particular form appears in the input” whereas the latter refers to “the number of distinct lexical items that can be substituted in a given slot in a construction” (N. Ellis et al., 2016, p. 52). In the present study, token frequency is most relevant, and is what I refer to when, for instance, general academic words are described as infrequent (see section 2.1.4).

Overall, the effects of frequency on language learning support usage-based approaches where the role of input is emphasized (N. Ellis, 2012). However, it is worth noting what Milton (2009) points out, namely that the effect of word frequency is not an absolute, but a tendency. He mentions word difficulty and learning burden as two factors that may weaken the effects of frequency (Milton, 2009). The learning burden refers to the effort involved in learning a word, and the notion is that the more familiar a word is for the learner, the lighter is the learning burden (Nation, 2022). Similarity with an L1 or other L2s contribute to this familiarity. However, from a usage-based perspective, similarity might also be a factor which impedes the acquisition of vocabulary in an L2. For instance, Verspoor and Schmitt (2013) highlight that words similar in form might be difficult to learn, as distinguishing them can be challenging. Problems may arise in the process of pairing form and meaning in cases where the features separating the similar forms are not very noticeable.

Along these lines, Laufer (1988) identified ten categories of what she calls “synforms”, defined as “words similar in form, phonological, graphic or morphological” (p. 117). She argues that such instances of lexical similarity can cause confusion among L2 learners of English, leading them to produce a similar, but wrong word. This taxonomy likely applies to all the basic skills, including writing, according to the author. Examples of synforms include words similar in form where only a vowel, suffix/prefix, or consonant separates them, such as *novel/novelty*, *curse/course*, and *propose/purpose*. From a usage-based perspective, experience with the language through input and output is considered essential for L2 language acquisition. However, in instances such as those elucidated above, differences in forms may be difficult to notice.

Thus, in certain cases, frequency alone may not be sufficient for the acquisition of particular lexis. If this was the only factor contributing to learning, “we would never get beyond the definite article” (N. Ellis, 2002, p. 178). This brings us to other issues such as awareness, noticing, and salience.

### **2.1.3 Salience and Noticing**

Mitchell et al. (2019) point out that there has been an interest in attention and awareness in second language learning research, primarily for two reasons. First, as a reaction to Krashen and his view of language learning as mainly implicit or incidental. According to Krashen (1989) and his Input Hypothesis, “comprehensible input is the essential environmental ingredient” (p. 440). He highlights reading as the most important source of comprehensible input for vocabulary acquisition. Second, the interest in attention and awareness can be attributed to their pertinence for language instruction (Mitchell et al., 2019).

On the one hand, attention is a cognitive process and includes the capacity of focusing on specific stimuli in the input, while disregarding others (VanPatten & Benati, 2010). On the other hand, awareness is often simply defined as consciousness. This relates to whether people are conscious of their actions, or more specifically in SLA, of their learning (VanPatten & Benati, 2010). These topics are connected to the issue of noticing, and all of them are “inextricably intertwined with salience” (Gass et al., 2018, p. 10). The term salience can be defined as the extent to which something stands out from the surroundings (VanPatten & Benati, 2010). Wulff and N. Ellis (2018) highlight the impact of salience on learnability and posit that “[s]alient items or features are more likely to be perceived, to be attended to, and are more likely to enter into subsequent cognitive processing and learning” (p. 43).

Moreover, Crossley et al. (2016) conducted a longitudinal study investigating the association between L2 input and output in relation to lexical development. Taking a usage-based approach, the authors examined aspects connected to the salience of words, such as meaningfulness or concreteness. The authors postulate that for L2 vocabulary acquisition, “salience may be a more robust indicator of acquisition than frequency” (Crossley et al., 2016, p. 705). Results from their study indicate that more salient words in the input are more likely to be produced by L2 learners in their output. The reverse is also true, that is, low salience might be a factor impeding learning (N. Ellis, 2012). However, it is worth noting that although Crossley et al. (2016) contrasts salience with frequency, the two concepts are not necessarily unrelated. In fact, salience may be affected by frequency. Both words that are noticeably frequent, but also infrequent, can be more salient (Loewen & Reinders, 2011).

Thus, words we encounter often and only rarely have the ability to stand out, and consequently catch our attention. This ability to stand out in the input refers to perceptual salience (Gass et al., 2018).

Along these lines, Schmidt (1995) asks whether L2 learning can take place without intention, attention, noticing, and understanding. His Noticing Hypothesis claims that “subliminal language learning is impossible, and that intake is what learners consciously notice” (Schmidt, 1990, p. 149). In other words, he maintains that attention and awareness are crucial for language learning irrespective of linguistic units. Crossley et al. (2014) contemplate on the importance of input, which is considered essential in usage-based approaches, and suggest that it is not sufficient by itself. The authors make the argument that acquisition is also dependent on other factors, and highlight noticing, among others (Crossley et al., 2014). These issues are essential in a usage-based approach to language acquisition, because “[w]hat is attended to becomes the focus of learning, and so attention controls the acquisition of language itself” (N. Ellis et al., 2016, p. 23).

Teachers can take advantage of this. By focusing the learners’ attention to particular words, the form-meaning connections may be strengthened (Verspoor & Schmitt, 2013). R. Ellis and Shintani (2014) highlight ways in which the input can be modified, making target vocabulary more salient and as a result, easier to notice. First, we may enrich the input with the words we want learners to acquire. Second, we can emphasize words in oral input by stressing them, or by highlighting them in written input, for instance through italicization or boldfacing. Third, text-elaboration, such as glossing, can increase the noticability of vocabulary. These strategies can be considered examples of constructed salience (Gass et al., 2018). While salience is one factor contributing to noticing, it may also occur through other means, such as using dictionaries, guessing from context, being explained a word or intentionally study it (Nation, 2022). The issues of salience and noticing therefore relate directly to the explicit and implicit dichotomy (Verspoor & Schmitt, 2013).

#### **2.1.4 The Implicit and Explicit Dichotomy**

There are three important distinctions in SLA research, which are implicit and explicit learning, knowledge and instruction (R. Ellis, 2009). These dichotomies can be discussed in isolation, or in relation to each other. DeKeyser (2017) underscores the importance of considering the relationship between these terms, and posits that it is important to establish what knowledge we want learners to acquire, and what implications this has for learning and instruction. The terms explicit and implicit knowledge may be defined as knowledge with or

without awareness, respectively (DeKeyser, 2017). Along similar lines, explicit learning might be defined as an intentional and conscious process, while implicit learning can be defined as learning where awareness and intentionality are absent (R. Ellis, 2008). On the one hand, the claim that incidental or implicit learning is possible, is not controversial in SLA (Ortega, 2009). However, on the other hand, what is not agreed on is the extent to which this learning can take place, or how efficient it is compared to explicit learning. An important question is therefore “[h]ow much implicit learning is possible and how much explicit learning is necessary?” (DeKeyser, 2017, p. 18). The fact that usage-based approaches consider frequency effects to be a determining factor in language learning suggests that learning is implicit (N. Ellis, 2002).

Nonetheless, the role of noticing and explicit instruction might still be important, also from a usage-based perspective. Linguistic items low in salience may require explicit instruction. This recognition is, however, not in opposition to a usage-based approach (N. Ellis, 2002). On the issue of attention, Römer (2023) postulates that language teachers need to focus learners’ attention to items that might be low in salience in the input, or otherwise challenging. In other words, within a usage-based model, “[language] learning proceeds by dynamic interactions between implicit and explicit processing” (Tyler & Ortega, 2018b, p. 318). Rather than viewing incidental and intentional vocabulary learning as polar opposites, it can be useful to view them in a complementary relationship (Hulstijn, 2001). Indeed, Schmitt (2008) argues that one requires the other. Incidental approaches are required to consolidate and enhance vocabulary taught explicitly, and intentional approaches may be necessary for vocabulary learned incidentally to become productive.

General academic vocabulary may often be considered items low in salience. Coxhead (2000) makes this very point, and exemplifies with the words *substitute*, *underlie*, *establish*, and *inherent*. She attributes their low salience in academic texts to the fact that they are not central, but rather supportive of, the topics in texts. Similar points have been made by others, such as Corson (1997) and Nation (2022), who attribute the difficulty of acquiring academic words to them being abstract, as well as their low frequency. This particular lexis may therefore be more difficult to acquire implicitly, or incidentally. Along similar lines, Hinkel (2003) suggests that to prepare learners for academic success, and expand their lexical repertoire, it might be insufficient with “mere exposure to academic text and reading” (p. 297). While reading academic texts is important, Cons (2012) argues that to improve the production of academic vocabulary, learners in secondary education also need more writing practice, as well as explicit instruction related to how academic words should be used in their

writing. Thus, explicit instruction may be required, in addition to extensive input and the ability to practice output in meaningful communication.

The effect of explicit vocabulary instruction on L2 writing was investigated in Lee's (2003) study of 65 intermediate ESL learners in a Canadian upper secondary school. In her study, explicit instruction was based on principles derived from psycholinguistics. Some examples include the teacher writing target vocabulary on the chalkboard, pronouncing it, and having the learners repeat, as well as explaining how vocabulary can be used in different contexts. In a pre-instruction writing task, only 13.19% of receptively known words were produced. However, in a post-instruction writing task, this increased to 63.62%. 23 days later, in a delayed writing task, this dropped to 55.46%. On the whole, this suggests that explicit instruction might be beneficial in order to elicit the written production of words that learners know receptively.

On a different note, we may consider how a usage-based approach to L2 learning fits within a communicative paradigm. According to N. Ellis (2002), usage-based and communicative approaches have in common the emphasis on exposure to input and practice of output. This point is underlined by Robinson and N. Ellis (2008):

A second language classroom providing learners with plentiful exposure to meaningful input, and opportunities to use the L2 while performing realistic communicative activities would be complementary, therefore, to the “input-driven” and “experiential” assumptions of usage-based learning. (p. 495)

Such an approach to language learning is therefore in line with a communicative paradigm, where the central aspect for all skills is to negotiate meaning (Skulstad, 2020). Indeed, usage-based approaches view contextually embedded participation and interaction as central to language learning, in which individuals communicate in order to achieve certain goals (N. Ellis et al., 2016). Overall, all usage-based models adhere to the tenet that communication is the main objective of language (Tyler, 2010). Another tenet is that meaning is central to language learning (Tyler & Ortega, 2018a). Thus, to summarize, a usage-based approach is in alignment with the communicative paradigm, considering the shared emphasis on the negotiation and conveyance of meaning in communication, as essential for second language learning and development.

Finally, Kemmer and Barlow (2000) underline that usage-based theories uphold that when studying language one should study actual, authentic usage. This methodological practice is one factor uniting usage-based researchers (Bybee & Beckner, 2015). Thus, corpora, which contain usage data, provide an ideal foundation for analyzing language when

taking a usage-based approach. This indicates a compatibility between usage-based approaches to SLA and corpus linguistics (Wulff, 2021).

## 2.2 Corpus Linguistics

In the current study, corpus linguistics is used as a method. In SLA research, as well as in language teaching, corpora have provided unique information about language use (Römer, 2013). Corpora can be defined as “principled collections of naturally occurring spoken, written, or multimodal data which are stored in electronic format for quantitative and qualitative analysis” (Dang, 2020, p. 288). These data consist of machine-readable texts and are often intended to represent certain kinds of text (Nesi, 2016). Further, software tools are commonly utilized, allowing the researcher to navigate and make sense of the large amount of data (McEnery & Hardie, 2012). Two common examples are tools that show frequency data and concordances, which exemplify quantitative and qualitative analyses, respectively. Both are important in corpus linguistics (McEnery & Hardie, 2012).

Utilizing a corpus or multiple corpora for the study of language is central in corpus linguistics. Studies in this field are interested in authentic language use. For this reason, corpus linguistics is often located in opposition to Chomskyan linguistics (Adolphs & Lin, 2011; Flowerdew, 2004). Nonetheless, the use of authentic language as data underpins the main argument for using a corpus, namely that we do not need to rely on the intuition of native speakers to investigate language use (Hunston, 2002). Corpus data have provided information about how academic words operate in authentic academic discourses, such as university lectures, scholarly articles, discipline textbooks and so on (e.g. Biber, 2006; Gardner & Davies, 2014). Moreover, in the field of vocabulary studies, Schmitt (2010) maintains that:

corpora have transformed the way we think about and research vocabulary. It is hard to imagine any area of vocabulary research into acquisition, processing, pedagogy, or assessment where the insights available from corpus analysis would not be valuable. In fact, it is probably not too extreme to say that most sound vocabulary research will have some corpus element. (p. 307)

However, all studies in corpus linguistics are not homogenous. We can distinguish between corpus-driven and corpus-based approaches (Biber, 2010). Research taking the first approach aims to discover novel linguistic constructs by analyzing corpus data inductively (Biber, 2010). The current study takes a corpus-based approach, however. This type of research

attempts to give an account of how already established linguistic constructs are used in various types of data (Biber, 2010). While this approach sees corpus linguistics as a method, a corpus-driven approach rejects this notion (McEnery & Hardie, 2012).

Other distinctions in the field of corpus linguistics are summarized in Table 1:

**Table 1**

*Overview of Distinctions in Corpus Linguistics (Adapted from McEnery and Hardie, 2012, p. 3)*

Mode of communication	Written, spoken, sign language, gestures
Data collection regime	Monitor corpora, Web as Corpus, sample corpora, opportunistic corpora
Encoding	Annotated, unannotated
Number of languages	Monolingual, bilingual, multilingual
Analysis approach	Total accountability, data selection

The approach taken in the current study uses language produced in the written form, in a single language, that is, English. While it could be classified as a sample corpus, there are practical considerations which put certain constraints on notions of representativeness and balance, which are important in sample corpora. Although the texts are not annotated in the corpus, metadata about the texts, that is, information about achieved grades, the tasks, and preparation material are stored separately. Finally, regarding total accountability versus data selection, I aim towards avoiding confirmation bias. Nonetheless, the corpus is small and specialized (see section 3.3.1), and only certain items will be profiled and analyzed (i.e. vocabulary from the AVL). Thus, it falls somewhere in the middle of this distinction.

### **2.2.1 General and Specialized Corpora**

Corpora can be divided into two main types: general and specialized. There are two trends in corpus linguistics, according to Koester (2022), namely that general corpora are getting larger, while the compilation of small, specialized corpora is increasing. General corpora often consist of millions of words and are commonly used as a reference for comparison with smaller corpora (Szudarski, 2018). Two frequently used general corpora are the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA). Such corpora often represent “language as a whole” (Cobb & Horst, 2015, p. 189). Based on data from large, general corpora, researchers have been able to identify which words occur more and less frequently in the English language. For instance, from the corpora mentioned above,



Nation (2017) developed the BNC/COCA lists containing words listed by frequency. Various ways of categorizing vocabulary based on their frequency in large, general corpora, have been suggested. Schmitt and Schmitt (2014) suggest the label “high-frequency” for the 3,000 most frequent word families, “low-frequency” for word families beyond the 9,000 frequency level, and “mid-frequency” as the word families in-between. The concept of word families is further discussed in section 2.2.2. Overall, corpus derived frequency information has been useful in language teaching, where the general idea is that the more frequent words are most important to learn first (Vilkaitė-Lozdienė & Schmitt, 2020). The notion of frequency is also emphasized in usage-based theories of L2 vocabulary acquisition (see section 2.1.2).

General corpora are useful for a wide range of purposes. However, the most important area of progress in corpus linguistics might be related to specialized corpora, providing data on language usage in specific registers (Reppen & Simpson-Vlach, 2020). Specialized corpora are defined by McEnery and Brookes (2022) as corpora created with the intention to represent a specific language variety or genre, often in a particular time or setting. Both specialized and general corpora have been useful for research into general academic vocabulary, particularly in the creation of word lists (see section 2.5.1 and 2.5.2).

One example of specialized corpora is a corpus consisting of learner language. This is called a learner corpus, which is what we can define the corpus compiled for the current study as. The use of this kind of corpora has established its own field of research, termed learner corpus research (LCR). This is an interdisciplinary field, methodologically similar to corpus linguistics, and theoretically similar to SLA (Paquot & Plonsky, 2017). Granger (2002) emphasizes that learner corpora may provide informative insights related to both SLA and Foreign Language (FL) teaching. Thus, one aspect of investigating learner language in corpus linguistics is to improve or inform language teaching in an L2. In particular, learner corpora have been used to analyze errors in learner language, or to compare it with native language (Granger, 2009). Furthermore, Gilquin et al. (2007) posit that the comparison of learner corpora with native corpora can provide insight into what linguistic items learners use extensively or to a limited extent, when writing in English for Academic Purposes (EAP). Along similar lines, Flowerdew (2015) points out that in EAP research, comparisons between learner corpora and academic word lists are common, in order to identify which items occur in learners’ written production. Thus, on the one hand, learner corpora may be compared to native corpora in order to describe what is specific to learners’ written production, as described by Gilquin et al. (2007). On the other hand, learner corpora may be compared to academic word lists to see the extent to which word list items are used by the learners

(Flowerdew, 2015). In the current study, I take the second approach, by comparing a small, specialized corpus of learner writing to the AVL by Gardner and Davies (2014).

### 2.2.2 Word Count Units

Furthermore, corpus linguistics used for the purpose of researching vocabulary requires us to define terminology related to what counts as words (Szudarski, 2018). Nation (2022) highlights five different counting units: tokens, types, lemmas, flemmas, and word families. If we count tokens, then every occurrence of the word is counted, regardless of whether the same word is repeated. Tokens and running words are used synonymously. Types, on the other hand, refers to different words, that is, the same words are not counted more than once. Thus, the sentence “it is what it is” consists of five tokens, but only three types, as *it* and *is* are repeated. Whether to count different words or running words is the most essential distinction to make, according to Nation (2022). If we count types, we would count *house* (singular) and *houses* (plural) as two different words. On the other hand, if we count lemmas, we would count the base word, its inflections, and its reduced forms, together (Nation, 2022). While only words belonging to the same part of speech (POS) would be included in a lemma, this is not the case for flemmas. If we use flemmas as the counting unit, we may count words of different POS, such as nouns and verbs, as the same word if the form of the base word is the same (Nation, 2022). To illustrate, consider the verb *view* and the noun *view*. These two words, and each of their inflections, would be grouped together as one flemma, but as two separate lemmas.

Moreover, if we count the word family, we include all semantically related word forms (Schmitt, 2010). This includes both inflections and derivations. Nonetheless, what should be considered the appropriate word count unit in vocabulary studies is not resolved and continues to be debated. When counting specialized, technical vocabulary, the appropriate counting unit might be the word type (Chung & Nation, 2003). However, some argue in favor of the lemma, as it offers more precision, while avoiding assumptions of learner knowledge (Gablasova & Brezina, 2021). Schmitt (2010) postulates that the lemma is the best unit given the easy process of lemmatizing words and the unambiguous process of interpreting them. Others, however, defend the use of word families (Laufer, 2021). Indeed, Schmitt (2010) acknowledges that word families might be “the best at capturing all of the word forms related to a concept” (Schmitt, 2010, p. 190). However, according to Nation (2016), types, lemmas and flemmas are different levels of Bauer and Nation’s (1993) word family scale. Thus, he maintains that the more useful discussion would be to decide the appropriate word family

level for certain purposes and learners (Nation, 2016). On the whole, however, Webb (2021) points out that for any purpose, there might be advantages and disadvantages associated with the use of any given counting unit. Deciding on what counting unit to use in vocabulary is an important task, as it will greatly impact the results (Schmitt & Schmitt, 2020). The counting unit in the word list used in the current study, the AVL, is the lemma (see section 2.5.2).

## **2.3 Vocabulary Knowledge**

While deciding on the proper word count unit is one important task when studying vocabulary, even after establishing this, it may still be challenging to define vocabulary knowledge. R. Ellis (2008) elucidates why this might be the case:

There is no agreed terminal stage for knowledge of a word. In fact, there is no such thing as a target-language lexicon because vocabulary, in contrast to grammar, constitutes an open system and individual native speakers vary enormously both in the size of their lexicon and in their DEPTH OF KNOWLEDGE of specific words. (p. 99)

This complexity notwithstanding, two concepts used to describe vocabulary knowledge are highlighted in the passage above: size and depth. This is one distinction developed to grapple with the complex mental lexicon (Schmitt, 2014). Size (or breadth) refers to the number of words known to a learner, whereas depth (or quality) refers to how well the learner knows these words (Schmitt, 2014). While vocabulary size is important, the depth of knowledge is also essential. It is crucial for learners to know vocabulary well, “to use them productively, appropriately, and fluently” (Schmitt, 2014, p. 942). This is of particular interest in the present study where productive use of general academic vocabulary is examined.

### **2.3.1 Receptive and Productive Knowledge**

Distinguishing between receptive and productive knowledge is one way to conceptualize vocabulary knowledge. Some use the terms passive and active instead of receptive and productive (e.g., Corson, 1997; Laufer, 2005). I use the latter terms in this study, however, when describing research, the terms preferred by the authors are used. While receptive knowledge refers to the ability to comprehend and retrieve the meaning of vocabulary through reading or listening, productive knowledge refers to the ability to retrieve and produce vocabulary in writing or speech (Nation, 2022; Schmitt, 2014). Receptive knowledge is considered to be larger in size than productive knowledge. In other words, “people know more words than they use” (Hirsh, 2010, p. 226).

To address this difference, Webb (2008) gave 83 EFL students in a Japanese university two translation tests, one L2-L1 meaning recall test to measure receptive knowledge and one L1-L2 form-recall test to measure productive knowledge. Results were scored sensitive and strict, where correct spelling was required for the latter but not the former scoring level. Webb (2008) found that the ratio of productive to receptive knowledge was 77% using strict scoring, and 93% using sensitive scoring. In other words, the results indicate that receptive vocabulary knowledge is larger in size than productive vocabulary knowledge. Further, by comparing the results from two tests, one of passive recall and one of active recall, Laufer (2005) found that the ratio between active and passive knowledge ranged from 16-35%. In other words, only 16-35% of the receptive vocabulary knowledge was produced in the active recall test. Similarly, Lee (2003) found that only 13.19% of recognized vocabulary was produced in a post-reading writing task (see section 2.1.4). However, it is worth mentioning that receptive tests often allow for guessing, something productive tests do not (Schmitt, 2014; Webb, 2008). This could also contribute to the gap.

Furthermore, Webb (2008) examined the impact of the frequency of the target words in the translation tests on receptive and productive knowledge, and the ratio between them. These differences were most significant when results were scored strict. He found that, as words became more infrequent, both receptive and productive scores decreased. At the same time, the ratio increased. This indicates that for less frequent words, learners know fewer words receptively, and might struggle more to retrieve productively the words they do know. However, this ratio of productive to receptive knowledge became greater for learners with the lowest receptive score, than for the learners with the highest receptive score. Webb (2008) therefore argues that his results indicate that “[l]earners who have a larger receptive vocabulary are likely to know more of those words productively than learners who have a small receptive vocabulary” (p. 91).

### **2.3.2 Vocabulary and Writing**

Knowledge of vocabulary is crucial to writing (Nation, 2022). While receptive knowledge is considered to precede, and to some extent indicate, productive knowledge, it does not automatically lead to correct productive use (Laufer, 1998; Laufer & Paribakht, 1998; Lee, 2003). When writing texts, several different aspects of vocabulary knowledge are required. These aspects are related to the form, meaning and use of a word. Nation (2022) provides an overview of what learners need to know when writing. This is presented in Table 2.

**Table 2**

*Aspects of Productive Vocabulary Knowledge Needed for Writing (Adapted from Nation, 2022, p. 54, with Inspiration from Coxhead, 2012, p. 138)*

<b>Form</b>	Written	How is the word written and spelled?
	Form and meaning	What word form can be used to express this meaning?
<b>Meaning</b>	Concept and referents	What items can the concept refer to?
	Associations	What other words could we use instead of this one?
<b>Use</b>	Grammatical functions	In what patterns must we use this word?
	Collocations	What words or types of words must we use with this one?
	Constraints on use (register, frequency ...)	Where, when, and how often can we use this word?

This overview illustrates the complexity of knowledge involved in writing. Productive knowledge of words is considered more difficult to acquire (Lin & Morrison, 2010), as well as more difficult to use, compared to receptive acquisition and use (Nation, 2022). Why this is the case continues to be debated, but the aspects of productive word knowledge necessary for writing highlighted in Table 2 are among the reasons.

N. Ellis and Beaton (1993) suggest that this difficulty of writing may be because productive learning requires more knowledge, such as spelling, or that learners have less practice in productive use compared to receptive use. Webb's (2008) study elucidates how the factor of spelling contributes to increasing the gap between receptive and productive word knowledge. In addition to spelling, Laufer (2005) highlights knowledge of register and collocations as aspects contributing to making productive use more difficult than receptive use. Furthermore, when writing, learners cannot utilize coping strategies to mitigate the lack of vocabulary knowledge which are available when reading, such as to skip unfamiliar words or guess the meaning of words based on context (Durrant, 2016). The difficulty of using productive vocabulary knowledge, compared to receptive, suggests that the former might be a more complete form of vocabulary knowledge. Indeed, according to Corson (1997), "[w]ords are only fully learned when they are available for active use. When words sit only in passive vocabulary, with their active use inhibited, then full word learning has not taken place" (p. 699).

Furthermore, there is a difference between production in vocabulary tests and actual use of words in written production of texts. Laufer (2005) notes that vocabulary elicited in

productive vocabulary tests, might not necessarily be used by learners in free written production. This is where vocabulary choice is up to the learners. For this reason, Laufer (1998) differentiates between two types of productive vocabulary knowledge “controlled active” and “free active”. While the first type would be elicited by a specific task or test, the latter relates to words used freely in written composition. Laufer (1998) found in her study, that while the size of the learners’ receptive and controlled active vocabulary size increased, they did not put these words to use in free active composition. She postulates that the learners’ “free active vocabulary reached a plateau beyond which it did not progress” (p. 266) where it had “fossilized” (p. 267). She attributes this to a strategy of convenient performance, where learners use as few words as possible to convey meaning, defined as “a risk-avoiding strategy, task simplification, or simply, taking the easy way out” (Laufer, 1998, p. 267). Reasons why learners use these strategies might be because they are not encouraged to take risks and use error-prone, advanced vocabulary. Teachers who are communicatively oriented might be satisfied as long as learners are able to convey meaning when writing (Laufer, 1998).

Along similar lines, Hasselgren (1994) investigated learners’ lexical choices in writing. She developed a taxonomy for lexical errors, where she investigated both the routes, but also the effects of wrong word choices when writing. Results showed that the effect of wrong word choice was most significant in terms of semantic dissonance.<sup>1</sup> This effect arises when a wrong word is produced and leads to wrongness in meaning, such as *distributes* instead of *contributes* or *went* instead of *walked* (Hasselgren, 1994). However, she discovered that collocational and stylistic/connotational dissonance combined were almost as significant, neither of which entails any wrongness of meaning. The first effect occurs when a word feels wrong, due to disharmony with surrounding words. The second effect arises when the style, formality or connotation is inappropriate (Hasselgren, 1994). In a partial replication of this study, Mahan (2013) found resembling results among Norwegian intermediate and advanced learners of English, arguing that her research “suggests that above all, Norwegian students of English are unaware of what types of register English words belong to” (Mahan, 2013, p. 68).

In other words, both Hasselgren (1994) and Mahan (2013) found the effect of wrong word choice to be error in style rather than error in meaning, both for advanced and intermediate learners of English. Hasselgren (1994) attributes this to the dependence on safe,

---

<sup>1</sup> Hasselgren (1994, p. 242) points out that this category, in addition to stylistic, collocational, and syntactic dissonance, was taken directly from Carter (1987). However, Carter (1987, p. 65) assigns their origin to Martin (1984).

general words termed lexical teddy bears. These teddy bears are general in meaning and familiar, that is, they are either learned at an early age, or influenced by the learner's first language (L1). She found this to be most apparent in Norwegian learners' avoidance of specific collocations often used by native speakers, and instead, their overuse of cores, which is one of the routes in her taxonomy. This is defined as words that are "widespread in usage, neutral in style or connotations, and high in frequency" (Hasselgren, 1994, p. 253). For general purposes, such general high-frequency vocabulary is indeed useful, however, it is often considered "the enemy of effective academic writing" (Granger & Larsson, 2021, p. 1).

These findings might indicate similar tendencies of risk-avoiding strategies, suggested by Laufer (1998). However, it might also suggest that writing texts, especially in educational contexts where a formal style is expected, is difficult. Indeed, research shows that students report finding productive skills more challenging than receptive skills, with regard to academic English. In a study of students in English medium instruction (EMI) classes at a Turkish university, Kamaşak et al. (2021) looked at what challenges related to academic language were most prominent. About two thirds of the participants were Turkish students, while one third were international. The self-reported measures showed that students found productive skills (i.e. speaking and writing) to be more challenging than receptive skills (i.e. listening and reading). Among the most salient challenges, was to write in a proper academic style (Kamaşak et al., 2021).

On the whole, producing vocabulary of appropriate style in writing can be difficult, especially in an L2. Indeed, in the L2, there might be a gap between the meaning or content we want to convey, and the vocabulary at our disposal, which might be inadequate (Laufer, 2013). This difficulty may be especially true for the productive skill of writing. Because this mode often consists of more academic and infrequent words, learners need to select their vocabulary in a more intentional and careful manner (Laufer, 2013). This resonates with Nagy and Townsend's (2012) point, namely that many of the characteristics of academic language might apply more to the written mode, where these characteristics typically appear to a greater extent than in speech. The fact that even learners in university tend to overuse core vocabulary in their writing (Hasselgren, 1994), and also report struggling with the written academic style (Kamaşak et al., 2021), highlights the necessity to focus further on general academic vocabulary and productive use of this lexis.

## 2.4 Academic vocabulary

While the importance of academic language has been emphasized frequently in later years, defining it in simple terms is difficult (Snow & Uccelli, 2009). Nonetheless, attempts to do so have been made. Nagy and Townsend (2012) define academic language as “the specialized language, both oral and written, of academic settings that facilitates communication and thinking about disciplinary content” (p. 92). The authors point out that there are six interdependent, functionally related characteristics that academic language has more of compared to everyday language. These are:

1. Latin and Greek vocabulary
2. Morphologically complex words
3. Nouns, adjectives, and prepositions
4. Grammatical metaphor, including nominalizations
5. Informational density
6. Abstractness (Nagy & Townsend, 2012, p. 93).

In other words, academic language is multifaceted and complex. However, according to Nagy and Townsend (2012), the most apparent component of academic language is possibly academic vocabulary.

This lexis can be defined in two ways, as domain- or discipline-specific academic vocabulary, on the one hand, or as general academic vocabulary, on the other hand (Baumann & Graves, 2010; Nagy & Townsend, 2012). The former may be defined as “content-specific terms and expressions found in content area textbooks and other technical writing” (Baumann & Graves, 2010, p. 6). This kind of vocabulary is often unique to particular academic disciplines (Nagy & Townsend, 2012), but not necessarily. Nation (2022) defines this lexis as technical rather than discipline- or domain-specific vocabulary. He underlines that the most important characteristic for this lexis is that it relates to the content of certain disciplines, in terms of meaning. In order to understand any particular domain, such as law or medicine, technical vocabulary is thus essential (Schmitt & Schmitt, 2020). However, it might be more difficult for teachers to help learners with technical vocabulary than academic vocabulary (Nation, 2022). The reason is because technical words require subject or discipline knowledge and should generally be learned along with the content of the field. In comparison, academic vocabulary, although specialized as well, is more detached from specific disciplines and subjects, and teachers may therefore be able to provide more help with this lexis (Nation, 2022). Moreover, general studies in upper secondary school in Norway is intended to prepare learners for higher education, rather than any specific area of study.



In the current study, the focus is on general academic vocabulary. This lexis is defined by Dang et al. (2017) as “words that have high frequency, wide range, and even distribution in a corpus representing materials from different academic subject areas” (p. 963). In other words, this lexis is not discipline or subject specific, but occurs in academic language across disciplines, and less in non-academic language (Nagy & Townsend, 2012). General academic words have been shown to cover 10-14% of the running words in different academic corpora (Coxhead, 2000; Gardner & Davies, 2014). Corpora of other texts contain a significantly smaller number of general academic vocabulary tokens, which is why this lexis is considered specialized (Coxhead, 2020). For instance, Gardner and Davies (2014) found that in the COCA and BNC corpora, 7-8% of tokens in the newspaper sections and 3.4% of tokens in the fiction sections, were made up of general academic vocabulary. Thus, in a sense, there is nothing particular about these words that make them inherently academic. Instead, what defines them as general academic vocabulary is that the words occur more often in academic texts than in other types of texts (Schmitt & Schmitt, 2020). Yet, others have emphasized that general academic words do have certain characteristics, such as a “formal written academic flavour” (Nation, 2022, p. 275). It has also been underlined that much of this lexis is Graeco-Latin (Corson, 1997; Coxhead, 2000; Nagy & Townsend, 2012; Nation, 2022).

This Graeco-Latin nature is one factor contributing to the difficulty to learn general academic vocabulary. Corson (1997) emphasizes that other reasons why these words are challenging to learn and use, are that they tend to be infrequent, as well as “non-concrete, low in imagery, ... and semantically opaque” (p. 696). These words are less attached to specific subjects and their content (Nation, 2022). Rather than being central to the topic, general academic words are supportive of it (Coxhead, 2000). As discussed in section 2.1.4, this might make them low in salience and difficult to notice in the input. Considering that many learners might not encounter these words outside of school as well (Corson, 1995, 1997; Olsson & Sylvén, 2015), it might not be surprising that Norwegian learners have been shown to struggle with English academic language demands in tertiary education (Busby, 2015; Hellekjær, 2005, 2008, 2009). Indeed, research has shown that Norwegian L2 learners of English in upper secondary school have a limited receptive understanding of general academic vocabulary (Skjelde & Coxhead, 2020). Similar results have been found in other Scandinavian countries (e.g. Warnby, 2023).

While words known productively through use should be understood receptively as well, the reverse is not necessarily true (Laufer, 1998). As outlined in section 2.3, productive vocabulary is considered smaller in size, and more difficult to acquire and use, than receptive

vocabulary. Thus, given the limited receptive knowledge of general academic vocabulary among upper secondary learners indicated by previous studies, one might expect their productive knowledge to be even more restricted. However, studies have relied on different definitions of this lexis. There is not consensus on what specific words should be defined as general academic vocabulary. As outlined in the introduction, this lexis is usually defined by their inclusion on a corpus-based word list (Schmitt & Schmitt, 2020). Variation in the corpora used, criteria for word selection, and word count unit have resulted in general academic word lists containing considerably different sets of words. In the next section, the notion of word lists is presented, and specific general academic vocabulary lists are discussed.

## **2.5 Word lists**

L2 learners of English cannot learn the entire lexicon in the English language. Nor is it likely that they will grow their vocabulary at the same rate as native speakers of English (Siyanova-Chanturia & Webb, 2016). For these reasons, researchers interested in vocabulary learning in an L2 have been interested in the development of words lists, particularly since the introduction of Zipf's (1949) law (Nation, 2016). This law applies to vocabulary and can be described as: "a pattern where a relatively small number of high-frequency items are extremely frequent, but then the frequency drops off exponentially, with the vast majority of items becoming relatively rare quite quickly" (Schmitt & Schmitt, 2020, p. 7). There are several implications of this law. While a small number of words occur frequently and cover a substantial proportion of any text, the majority of words occur infrequently. Indeed, in lists based on frequency, approximately half of the words tend to only occur one time (Nation, 2016).

As mentioned in section 2.2.1, information about word frequency has informed decisions on what words are most important to focus on in language teaching. Along similar lines, the premise that word lists are based on is that certain words deserve more attention than others because of their usefulness (Durrant, 2016). The construction and validation of word lists for L2 learners have primarily been based on data from corpora (Dang, 2020).

One influential word list is the General Service List (GSL) developed by Michael West (1953). The list contains around 2,000 of the most common word families based on a corpus consisting of 5 million English words. It is considered an important and influential high-frequency word list (Nation, 2022), and the list provides words of general usefulness to L2 learners of English. However, it is not based only on frequency, nor other commonly used

objective criteria such as range or dispersion. Nation (2016) elaborates on the criteria behind the selection of the words on the GSL, and observes that subjective criteria were used as well, such as ease or difficulty of learning, necessity, and others. Using subjective criteria in combination with objective ones can have certain benefits (Dang, 2020). However, the fact that the GSL is based only on written text has been criticized (Nation & Meara, 2020). It is also considerably dated. Nevertheless, the words on the GSL have been shown to cover 70-90% of running words in various types of texts (Dang et al., 2022; Gilner, 2011).

With the principles and premises behind such a word list in mind, researchers have looked toward a similar list for general academic words. The important question in such an endeavor is how to discover groups of words similar enough to be represented by one list of words, and simultaneously apply to several different learners (Durrant et al., 2022). Indeed, Eldridge (2008) maintains that “[t]he notion of academia is a broad one, and it cannot easily sustain a generic word list” (p. 112). Nonetheless, attempts have been made to identify in advance which words are useful in English for general academic purposes. While several lists have been created, I will highlight the two that have been most influential.

### **2.5.1 The Academic Word List (AWL)**

Building on the GSL, Coxhead (2000) developed the AWL, a word list of general academic vocabulary. The word list is based on an academic corpus consisting of 3.5 million tokens, composed of four subcorpora, each of which is divided into seven subjects, leading to 28 different subject areas in total.<sup>2</sup> The AWL consists of 570 word families. This count unit was based on Level 6 of Bauer and Nation’s (1993) word family scale and is briefly defined as “stem plus all closely related affixed forms” (Coxhead, 2000, p. 218). The list covers about 10% of the running words in academic texts (Coxhead, 2000). It replaced the University Word List (UWL) by Xue and Nation (1984). Unlike the UWL, “the AWL was developed using a written academic corpus, validated on a second academic corpus, and contrasted with general English in another corpus” (Coxhead, 2016b, p. 181).

Three criteria guided the selection of words to include in the AWL. These criteria were range, frequency, and uniformity. Coxhead (2011) summarizes them in the following way, “[t]he word families had to occur 100 times or more in each of the four disciplines of the corpus (frequency), in 15 or more of the subject areas (range), and over 10 times in the four disciplines (uniformity)” (p. 356). In addition to these criteria, she excluded West’s (1953)

---

<sup>2</sup> See Table 2 in Coxhead (2000, p. 220) for a detailed overview of the composition.

GSL words. Thus, the AWL was built on top of the GSL. This was done based on an assumption that L2 learners of English learn high-frequency words of general usefulness before learning more specialized lexis, such as academic words (Schmitt & Schmitt, 2020). The exclusion of GSL items is one aspect of the AWL which has received criticism (Eldridge, 2008; Gardner & Davies, 2014; Paquot, 2010).

Nation (2016) points out that assuming learners have knowledge of the high frequency vocabulary in English may be inappropriate for L2 learners of English:

For learners in non-English speaking countries going on to do university study largely in their first language but also with an English requirement, the high frequency words may still be poorly known and thus the *Academic Word List* may be too big a step at least initially. (p. 11, emphasis in original)

However, one benefit of excluding general service words in the AWL is that both teachers and learners avoid repeating known words during language instruction (Dang, 2020). Yet, the AWL also contains words considered high-frequency words today (Gardner & Davies, 2014). In fact, Cobb (2010) found that when compared to the frequency levels in the BNC, 280 items from the AWL were located within the first two 1,000 levels. For these reasons, some have suggested that testing AWL knowledge among language learners might instead test their knowledge of general, rather than specialized words (Masrai & Milton, 2018).

Moreover, many GSL words might actually be general academic words (Nagy & Townsend, 2012). Due to the AWL's exclusion of these items in its word selection, such vocabulary is therefore not considered general academic words in Coxhead's (2000) list. In addition, while this study does not focus on multi-word units, it is still worth mentioning that GSL words might also be part of academic collocations (Durrant, 2009). In other words, there might be several problematic aspects related to the AWL's relationship to the GSL. These insights suggest the need for a new general academic vocabulary list for L2 learners of English.

### **2.5.2 The Academic Vocabulary List (AVL)**

A more recent general academic word list is the AVL by Gardner and Davies (2014). The AVL consists of 3,015 lemmas and is extracted from the academic sub-corpus of the COCA.<sup>3</sup> This sub-corpus is 120 million words in size, that is 34 times larger than the corpus Coxhead's

---

<sup>3</sup> Durrant (2016, p. 53) detected one word which occurred twice on the list. Thus, the AVL consists of 3,014 unique lemmas.

(2000) AWL is based on. The composition of the academic section in the COCA consists of nine disciplines.<sup>4</sup> This is one strength with the AVL, because general academic word lists should draw on corpora containing texts from a variety of disciplines (Dang, 2020). However, the majority of the words are made up from academic journals (85 million of the 120 million words). This might reduce the diversity of the source corpus to some extent.

Four criteria guided the selection of which words (lemmas) to include in the AVL: ratio, range, dispersion, and discipline measure (Gardner & Davies, 2014). First, the ratio for word-selection was set at 1.5. In other words, for a word to be included in the AVL, it had to occur 50% more frequently in the academic section of the COCA, than in the non-academic section. This is why the AVL is considered a corpus-comparison approach (Nation et al., 2016). Second, the range value was set to seven disciplines at 20%. Consequently, for a word to be included in the AVL, the expected frequency of that word had to exceed a threshold of 20% in at least seven out of the nine disciplines. Third, a dispersion score was set at 0.80. Dispersion in corpus linguistics “refers to the evenness with which an element is distributed throughout a corpus (Gries, 2022, p. 170). This means that for a word to be included in the AVL, its frequency had to be spread out across the corpus somewhat evenly. Fourth, a discipline measure was set, excluding any word that exceeded the expected frequency with more than three times, in any of the disciplines. While the first criterion was designed to exclude general high-frequency words, the other three criteria were designed to exclude technical and discipline-specific words (Gardner & Davies, 2014).

Besides being based on a significantly larger corpus, the AVL is also more updated, for the obvious reason that it was developed fourteen years later than the AWL. Additionally, the relationship with the GSL makes the AWL even more dated, considering West developed his list in 1953. One benefit, however, with the AWL is that it is divided into ten sub-lists, which makes it more useful for pedagogical purposes. The AVL, on the other hand, might be too large to readily make use of by teachers and students alike. For this reason, Webb and Nation (2017) suggest that this word list is more useful for researchers than teachers. Overall, however, Szudarski (2018) maintains that the AVL is likely to be widely used both in research and pedagogy related to vocabulary, because of the rigorous methodology behind the creation of the word list.

However, one of the most significant differences between the two general academic word lists is what counts as words. In contrast to the AWL, Gardner and Davies’ (2014) AVL

---

<sup>4</sup> See Gardner and Davies (2014, p. 314) for a detailed overview of the composition.

counts lemmas instead of word families. The different word count units were discussed in section 2.2.2. In addition to the exclusion of West's (1953) high-frequency words, the use of word families was among the most problematic aspects of Coxhead's (2000) AWL, according to Gardner and Davies (2014). Nonetheless, the authors developed a version of the list where the counting unit is word families, to allow for comparisons with the AWL. Table 3 compares the twenty most frequent word families in both lists, with a third column showing the twenty most frequent lemmas in the AVL, as well. It should be noted that Coxhead's (2000) AWL is organized alphabetically within each sublist, which is based on frequency, whereas the AVL is organized by frequency only.

**Table 3**

*Top 20 AWL Word Families, AVL Word Families, and AVL Lemmas*

AWL (families) 1-20		AVL (families) 1-20		AVL (lemmas) 1-20	
Analyse	Constitute	Study	However	Study (n)	Important (j)
Approach	Context	Develop	Increase	Group (n)	Process (n)
Area	Contract	Group	Experience	System (n)	Use (n)
Assess	Create	System	Level	Social (j)	Development (n)
Assume	Data	Relate	Process	Provide (v)	Data (n)
Authority	Define	Research	Culture	However (r)	Information (n)
Available	Derive	Social	History	Research (n)	Effect (n)
Benefit	Distribute	Result	Active	Level (n)	Change (n)
Concept	Economy	Use	Support	Result (n)	Table (n)
Consist	Environment	Provide	Individual	Include (v)	Policy (n)

*Note.* n = noun, v = verb, j = adjective, r = adverb.

The AVL and AWL might be said to represent two different approaches to the development of word lists (Coxhead, 2018). The AWL represents a common core approach, where previous vocabulary knowledge is taken into consideration, whereas the AVL represents a specialized approach, in which prior knowledge is not taken into account (Coxhead, 2018). In the creation of the AVL, prior knowledge of high-frequency general service words were not considered. Dang (2020) underlines a weakness associated with this approach, namely that already known words might be at risk of being repeated during language instruction. These different approaches are one of the reasons why the words in the two lists differ in significant ways. In a comparison between the two lists, Hartshorn and Hart (2016) found that 31.05% of the words overlap, 26.87% of AVL words are GSL words, whereas 42.07% of AVL words are neither found in the AWL, nor in the GSL.

Moreover, Coxhead (2020) also contrasts the AVL's common core approach with the AVL's corpus comparison approach. While she used the GSL to exclude high-frequency words, Gardner and Davies (2014) compared corpora to achieve a similar goal, as explained by the first criterion. Such a comparison was used to identify the vocabulary appearing more frequently in academic texts than in non-academic texts in English (Coxhead, 2020). However, one downside of this approach is that certain words can be included that do not appear academic. Because of this, Nation et al (2016) characterize the AVL as "only partially successful" (p. 150) and suggest that considering prior vocabulary knowledge might be necessary, while pointing to alternatives to the GSL. Gardner and Davies (2016) acknowledge that their list is not perfect, as a result of using quantitative statistics, but decided to leave it the way it was.

Furthermore, when applying the 570 AVL word families to the academic sections of the BNC and COCA, these words provide 13.7% and 13.8% coverage, respectively (Gardner & Davies, 2014). Each sub-section of the large general corpora contain approximately 32 and 120 million tokens. For comparison, Coxhead's AVL covers only about 7% of the academic sections in the same corpora (Gardner & Davies, 2014). Along similar lines, when applied to specialized corpora of learner language, the two word lists have been shown to provide different coverage. For instance, Olsson (2015) found that the AVL was better suited for detecting progress in academic vocabulary production in written texts by Swedish upper secondary learners, because it covered more words than the AVL. For the purpose of detecting progress in written academic vocabulary use, the AVL was therefore considered most useful. In Olsson's (2015) words, the AVL's higher coverage "contributes to a more detailed description of students' academic vocabulary use" (p. 68). Thus, compared to the AVL, the AVL has higher coverage in both academic corpora and corpora of learner language. Gardner and Davies (2014) argue that their list is "the most current, accurate, and comprehensive list of core academic vocabulary in existence today" (p. 325). Yet, Durrant (2016) underlines a weakness with the AVL, namely that it is largely concerned with receptive needs. This weakness is also directed towards word lists in general, where there is a tendency to focus on receptive vocabulary (Durrant, 2016). He maintains that it is of importance to focus on this lexis productively.

In order to evaluate the usefulness of the AVL for productive purposes, Durrant (2016) used the word list to profile successful university student writing, represented by the British Academic Written English (BAWE) corpus. He found that only 427 AVL items reached a threshold of 12 occurrences per million words in more than 90% of the disciplines in the

corpus.<sup>5</sup> He therefore suggests that these words can be considered important for students irrespective of discipline. He draws the conclusion that while a core productive academic vocabulary exists, the size is smaller than what the AVL suggest, with its 3,015 lemmas (Durrant, 2016). Similar indications are provided by Therova (2023). She found that only between 3.57% and 9% of AVL types were produced, on average, by at least half of the 193 L2 university students in her study, across four genres (exposition and discussion essays, problem questions, and research reports).

These findings are corroborated by Malmström et al. (2018) who used the same corpus as Durrant (2016). Attempting to follow the methodological choices of Gardner and Davies (2014), the authors compared the BAWE corpus to two reference corpora of student newspapers and blog posts, resulting in the Productive Academic Vocabulary List (PAVL). This list contains 591 word forms (or 474 lemmas), significantly smaller in size than the AVL, but similar to the findings in Durrant's (2016) study. Moreover, only 381 words forms (64%) on this list are AVL words. The authors therefore suggest that this portion of words may be considered a core academic vocabulary, as these are important for both receptive and productive purposes. In addition, the fact that around one third of the PAVL items are not on the AVL, suggests that there is a difference between students and published writers in terms of the academic words they use. Moreover, the authors found that the unique PAVL words, the words appearing on both the PAVL and the AVL, and those unique to the AVL, differ significantly in terms of their frequency in the COCA. These three groups of words have a median frequency in the COCA of 16,092, 9,576, and 1,539 respectively. In other words, published writers (represented by the AVL) use more low-frequency academic vocabulary, whereas student writers (represented by the PAVL) use more high-frequency academic vocabulary. Because of this, the authors suggest that, compared to published writers, "students have a distinct academic vocabulary, but it is not *very* academic" (Malmström et al., 2018, p. 36, emphasis in original).

Gardner and Davies (2016) responded to Durrant's (2016) criticism by strongly objecting to the notion that university students' writing represents disciplinary writing. This objection applies to Malmström et al. (2018) as well, considering the authors utilized the same BAWE corpus. However, the authors posit that since the student texts comprising the 6.5 million token corpus all received high grades, the corpus might be regarded as "a potential 'target' or 'norm' for students' English academic writing" (Malmström et al., 2018, p. 31).

---

<sup>5</sup> See table 11 in Durrant (2016) for an overview of the 427 AVL items.



Thus, they acknowledge that there is a difference between published academic writers and students, something Durrant (2016) also does. In addition, Gardner and Davies (2016) argue that Durrant (2016) understates the fact that the AVL in fact provided a significant coverage of the BAWE corpus. Indeed, of all the studies described further in section 2.6.1, the proportion of general academic vocabulary from the AVL is among the highest in Durrant's (2016) study. Finally, Gardner and Davies (2016) maintain that academic vocabulary can still be important to focus on, despite writers' limited use of this lexis. Thus, AVL words can be useful even though learners might struggle to produce them.

Furthermore, the notion of a core general academic vocabulary has been questioned from other perspectives, such as that of semantic disparity. For instance, Hyland and Tse (2007) state that "[i]t is by no means certain that there is a single literacy which university students need to acquire to participate in academic environments" (p. 236). To question this, they conducted a corpus study where they looked at items from the AWL, and examined their frequency, range, preferred meaning, and collocations. Their study showed that AWL items occur more often in certain disciplines while also taking on different meanings depending on the discipline. The authors argue that focusing on technical or specialized vocabulary specific to learners' needs, is more useful. Nagy and Townsend (2012) postulates that this polysemy in general academic words underlines the importance of teaching academic words in context. Moreover, responding to this criticism, Eldridge (2008) maintains that general academic word lists can still be of practical use, arguing that one of the purposes in research "is to find similarities and generalities that will facilitate instruction in an imperfect world" (p. 111).

Notwithstanding the arguments against a core academic vocabulary, only the AWL was investigated in Hyland and Tse's (2007) criticism of the usefulness of general academic word lists, as the AVL had not yet been developed. As mentioned above, the selection of items in the AVL was guided by a different set of criteria than in the AWL, where range, dispersion, and discipline measure were specific criteria designed to avoid discipline-specific vocabulary. In addition, part of Hyland and Tse's (2007) criticism against the AWL is connected to the use of word families, in that certain members of a word family are preferred in specific disciplines. The use of lemmas in the AVL mitigates this limitation to some extent. Nonetheless, the issue of semantic disparity is not absent when using lemmas. Both polysemous words and homonyms may be grouped together under the same lemma, although to a lesser extent than for word families (Gablasova & Brezina, 2021). Moreover, the AVL and how well it represents a core academic vocabulary was challenged by Malmström et al. (2018), as well as Durrant (2016), suggesting that the AVL in its entirety might be too

extensive, at least for productive purposes. On the whole, Durrant et al. (2022) describe the question of having a stable core vocabulary which can be made into one single word list as “an important, and currently unresolved, debate” (p. 76).

Despite the different ways in which the AVL has been challenged, this word list is utilized in the present study. Thus, general academic vocabulary is defined and operationalized as the words on the AVL. The reason why this list was chosen is because of the strict methodological criteria guiding the word selection, the use of a large and updated corpus, as well as the high coverage it has shown to provide compared to the AWL, both in corpora containing written academic texts and written learner texts (Gardner & Davies, 2014; Olsson, 2015).

## **2.6 Previous Research on Productive General Academic Vocabulary**

There has been a limited focus on L2 learners of English and their productive ability to use general academic vocabulary, compared to receptive abilities. This is problematic, as the use of general academic vocabulary in student’s written production of texts is necessary for academic success (Coxhead & Byrd, 2007). Moreover, the most common form of assessment in tertiary education is writing (Coxhead, 2012). For this reason, Lillis and Scott (2007) consider writing at university a high-stakes activity, arguing that insufficient writing skills might lead to students failing. The increasing role of English in tertiary education (e.g. Ljosland, 2007; Mauranen et al., 2010), suggests that L2 learners might be expected to produce general academic words in a variety of situations. Further, within academia, the majority of publications are written in English, mostly by non-native English speakers (Mauranen et al., 2016). This is true in Norway, as well, although the extent differs between disciplines (Ljosland, 2007). Thus, producing written English texts that are academic in nature might be required at several stages of tertiary education and in academia in general.

Therefore, because productive use of general academic vocabulary is also necessary, monitoring learners’ productive knowledge of this lexis is useful (Nation, 2022). There are various ways of measuring productive vocabulary knowledge, and the quality or sophistication of this production, which does not include testing. Lexical sophistication measures usually examine the extent to which a text consists of low frequency vocabulary (Nation, 2022). One commonly used example is the Lexical Frequency Profile (LFP) by Laufer and Nation (1995). Studies into lexical sophistication typically quantify the amount of vocabulary that is present on low-frequency word lists (Durrant et al., 2022). However, word

lists of general academic vocabulary have also been used in studies of lexical sophistication. Indeed, one approach is to measure “the presence of academic vocabulary” (Malmström et al., 2016, p. 53). This is the approach taken in the current study, and is a type of vocabulary profiling, where the proportion or coverage provided by academic words in a text is explored.

One limitation of such an approach is the same as its purpose, namely that vocabulary is given a binary status of whether it is on a specific word list, or not (Durrant et al., 2022). Schmitt (2010) argues that profiling tools may be insufficient, and that a measure of appropriacy of vocabulary in free writing composition would provide better insights. He suggests that it matters less what specific vocabulary is produced, and instead, whether it is used properly in a given context (Schmitt, 2010). However, profiling texts to measure the use of general academic vocabulary is one way to determine if students’ use appropriate vocabulary, given the significant role of academic vocabulary in characterizing academic language (Nagy & Townsend, 2012). Thus, as long as the profiled texts are expected to have a formal, academic language, I would maintain that general academic vocabulary use indicates appropriacy, in addition to sophistication, to some extent.

In the following sections, research on productive general academic vocabulary use in written texts, is presented.

### **2.6.1 General Academic Vocabulary Production in Written Texts**

Olsson (2015) examined general academic vocabulary coverage in 525 essays written by 230 L2 English learners throughout three years in upper secondary education in Sweden, elicited by four different tasks. The first task was written during the first year, the second and third task were written during the second year, and the fourth task was written in the final year. In addition to assignment task (1-4), one other variable was used, namely instruction method, divided into content and language integrated learning (CLIL) and non-CLIL. Although she profiled the learners’ texts using both the AWL and AVL, only results where the latter was used, is reported here. The mean AVL vocabulary coverage for the CLIL students was 7.37% (SD = 2.51), 9.30% (SD = 3.26), 8.47% (SD = 3.60), and 12.14% (SD = 3.51) in text 1-4, respectively. For the non-CLIL students, the average AVL vocabulary coverage was 5.85% (SD = 2.38), 7.93% (SD = 3.33), 6.65% (SD = 2.86), and 10.86% (SD = 3.59) in text 1-4, respectively. The productive use of AVL words was higher for the learners in the final year compared to the first year, irrespective of whether the instruction method was CLIL or non-CLIL. This indicates progress in AVL vocabulary use throughout upper secondary school.

Furthermore, Csomay and Prades (2018) investigated AVL vocabulary use in essays written by ESL university students at two instruction levels, one lower-level course, and one upper-level course. In addition to instruction level, the authors examined AVL vocabulary use across text types (response paper, comparative analysis, exploratory synthesis, argument synthesis, rhetorical analysis or editorial) and two drafts. The authors emphasize that the students had not received any explicit instruction of general academic vocabulary, only incidental learning from lectures and readings, as well as discussions in class (Csomay & Prades, 2018). The number of variables examined (instruction level, text type and draft) resulted in 16 different mean AVL vocabulary coverage values.<sup>6</sup> For text type and draft, the mean coverage ranged from 5.98% (SD = 1.68) to 12.06% (SD = 2.22). Descriptive statistics for draft and instruction level ranged from 8.57% (SD = 3.46) to 10.26% (SD = 2.65). Interestingly, students in the higher instruction level (upper-level) did not produce more general academic words than those in a lower instruction level (lower-level). The average AVL vocabulary production was 8.60% for students in the upper-level course and 10.25% for students in the lower-level course. On the whole, while draft did not have a strong effect on AVL vocabulary use, text type had a significant impact.

Malmström et al. (2016) investigated the coverage of AVL items in written texts produced by first and second year master of science students in a Swedish university, across four disciplines. The participants' degree programmes used EMI, and while some were international students, none of them had English as their L1. Their Master of Science corpus (MSc) consisted of 80 texts and approximately 720,000 tokens. Results showed that, overall, 19.3% of the tokens in the texts from their corpus were made up of academic vocabulary from the AVL. The high coverage of academic words is attributed to the use of the AVL instead of the AWL. The authors point to Gardner and Davies (2014) who found similar tendencies when applying the different word lists to the COCA and BNC corpora, where AVL vocabulary coverage was about twice of AWL vocabulary coverage (see section 2.5.2). In addition, Malmström et al. (2016) established that 70.5% of the AVL tokens were covered by the first 500 lemmas on the AVL word list. The authors point to a comparable proportion in the COCA, where the 500 most frequent AVL items make up 74% of the AVL tokens in the academic section. Finally, the authors compared the production of AVL tokens between first and second year students, as well as between home and international students. These

---

<sup>6</sup> See Table 10 and Table 13 in Csomay and Prades (2018) for the complete overview.

differences were low, ranging from 18.9% to 19.9% for the home and international students, and from 19.0% to 19.5% for students in the first and second master's year, respectively.

Csomay (2020) investigated academic vocabulary use in written assignment papers by Georgian EFL university students. She compiled a corpus of student papers written in seven different courses, from twelve different assignment types, resulting in approximately 1.9 million tokens. Besides profiling the student papers for academic vocabulary, through the use of the AVL, her study looked at how academic vocabulary use was affected by instruction level and assignment type. Similar to Csomay and Prades' (2018) study, these students did not receive any explicit teaching of academic vocabulary. Results showed that productive use of AVL vocabulary in student writing averaged 9.94% (SD = 2.80) for all courses. Moreover, findings indicate that this use did not differ significantly between courses at the same level. However, AVL use between instruction levels differed significantly. Papers in the two courses from year 1 had an average AVL vocabulary coverage of 7.57% (SD = 2.32) and 8.14% (SD = 1.71), respectively. Papers written in the two courses in year 3 had an average AVL vocabulary coverage of 11.96% (SD = 2.40) and 12.04% (SD = 2.90). This indicates "that students utilize incrementally more academic vocabulary as they progress in their studies" (Csomay, 2020, p. 22). In addition, findings indicate that assignment type had an impact on AVL use.

Further, the author underlines that one would often assume that the underuse of AVL words implies an overuse of general vocabulary. This could be what Hasselgren (1994) called lexical teddy bears. However, as Csomay (2020) maintains, a limited reliance on AVL words in written production may instead be the result of using specialized vocabulary which is domain-specific, indicating "a more sophisticated discussion of a given topic" (p. 24). Finally, two limitations with this study are emphasized by the author, namely the focus on single-word units and the lack of an examination of how AVL words are used. The first limitation applies to all of the studies reported in this section, and the present study as well. This is addressed further in section 6.3.2. On the second limitation, the author highlights looking at correct versus incorrect usage as an example which might provide a more in-depth and accurate analysis of written AVL production. Csomay (2020) underscores that this is difficult to accomplish in quantitative studies. However, in the current study where a small, specialized corpus is compiled, such analysis is more manageable.

Finally, Durrant (2016) profiled the BAWE corpus of successful student writing. Results showed an AVL vocabulary coverage of 16.82% (SD = 4.87) for all tokens, or 33.82% (SD = 8.40) for lexical tokens only. He argues that reporting the coverage of only

lexical words are more appropriate, given that the AVL word list itself only contains such words. Moreover, his results showed that approximately 10% of the AVL (313 words) provided two-thirds of the overall AVL lexical token coverage. He writes that, similar to other word lists, “the frequency of AVL items is highly skewed, and the majority of coverage is therefore achieved by a relatively small number of items” (Durrant, 2016, p. 56). Further, he examined variation based on discipline, level of study, and text genre. While there was only modest variation based on text genre, there was wider variation based on discipline. For instance, the mean AVL vocabulary coverage of lexical tokens in Economics texts was 39.59% (SD = 5.61), whereas in Classics texts, it was 21.39% (SD = 5.88). Based on level of study, the mean AVL vocabulary coverage of lexical tokens for level 1 and 4 was 29.19% (SD = 8.19) and 34.09% (SD = 8.93), respectively. Level 1-3 was each year of a British undergraduate degree, while level 4 was master’s level. This suggests that students develop their productive AVL use to some extent as they progress in their studies. Finally, it is important to keep in mind that, in contrast to the previously discussed studies in this section, the BAWE corpus utilized by Durrant (2016) does not consist of texts written by learners or students with English as their L2, specifically.

### **2.6.2 Factors Impacting General Academic Vocabulary Use: Source Texts, Topic, and Task**

Coxhead and Byrd (2007) highlight that there is a relationship between academic writing and reading of academic texts. The connection between the two skills relate to how reading is used for writing purposes. Part of the process of writing will usually consist of reading texts to take in data or ideas, but also to find appropriate language (Coxhead & Byrd, 2007). Hirvela (2016) also considers these two skills to be related. He points out that “in academic contexts, students are not often asked to write without some kind of stimulus to input, usually in the form of reading materials (i.e. source texts)” (p. 127). Although the reading and writing processes among learners in upper secondary education are likely to differ from what is described by Coxhead and Byrd (2007), this relationship is nonetheless manifest in the present study. Every learner was provided with preparation material that they had to read, prior to taking the written English mock exam, and the tasks consisted of texts of various lengths. Some of these texts, the learners had to respond to or interact with, others, they had to summarize to a given audience (see Table 5).

The impact of reading source texts has been investigated by Plakans and Gebiril (2012). The authors found results indicating that L2 university students used source texts to

acquire knowledge on a topic and develop opinions, but also for language support. This includes using source texts as a resource providing vocabulary, correct spelling, as well as technical terminology. Likewise, Leki and Carson (1997) found that some ESL university students experienced source texts as a form of scaffolding, which lessened the burden of having to find appropriate vocabulary for written assignments. Similarly, a case study conducted by Therova (2021) suggests that reading sources in the process of writing can lead to the acquisition and subsequent production of new AVL words.

Along these lines, Milton (2001) found that both Native English speakers and EFL students rely on words found in essay prompts in their written texts. However, the EFL students did this to a greater extent, and in particular, the least proficient ones. This link was investigated by Therova (2022), using the AVL to profile two essays written by 193 L2 university students. She compared the AVL tokens in these essays to the presence of AVL tokens in the assignment briefs. Results showed that, on average, 14.36% and 22.86% of the AVL tokens produced in essay 1 and 2, respectively, were from the assignment brief. The higher integration of AVL items from the assignment brief in essay 2, compared to essay 1, is attributed to the overall higher number of AVL items in the assignment brief for the second essay. First of all, these results suggest that students rely on the general academic vocabulary present in the assignment brief. Second, the results indicate that students' written texts contain a greater proportion of academic words when the assignment briefs also consist of a higher number of academic words (Therova, 2022). Thus, the assignment brief might not only serve an instructional purpose, providing information about what is expected in a writing assignment. It might also function as a source of appropriate vocabulary, for learners to include in their written texts (Therova, 2022). The author argues that teachers must be mindful of the way prompts are articulated, as well as to include vocabulary that they want learners to produce.

A related issue is the effect of topic on vocabulary. As Nation (2022) points out, the type of vocabulary which is used is influenced by "the general topic of the text" (p. 304). This was also investigated by Therova (2022), where she compared pairs of subcorpora differing based on assignment topic. She found that topic had an impact on the written AVL vocabulary production among students. In other words, that certain topics may elicit general academic vocabulary production to a greater extent, than others. Moreover, in Therova's (2021) case study, some of the AVL items produced most frequently by her participant, were related to the assignment topic.

However, others, such as Csomay and Prades (2018) suggest that genre or text-type influence AVL vocabulary production to a greater extent than topic, among students. Indeed, different tasks or assignments can influence the lexical choices made by learners, or “the kind of vocabulary that is elicited” (Read, 2000, p. 199). Research has shown this to be the case for general academic vocabulary, as well. Olsson (2015) observed this effect when she compared general academic vocabulary production in texts from four different assignments. Results from her study indicate that the extent to which academic words are elicited might vary based on task (Olsson, 2015).

### **2.6.3 General Academic Vocabulary Use and Assessment of Writing**

We know less about the relationship between vocabulary knowledge and productive skills, compared to receptive skills (Qian & Lin, 2020). Further, we know the least about this relationship for the productive skill of writing (Nation, 2022). Nonetheless, empirical studies have shown that vocabulary knowledge correlates with writing performance (Milton et al., 2010; Stæhr, 2008). Along these lines, Nation (2022) states that it is clear that “vocabulary plays a significant role in the assessment of the quality of written work” (p. 227). Besides profiling texts for general academic vocabulary use, studies have also examined the impact of the productive use of this lexis and assessment.

Csomay and Prades (2018) conducted a Pearson correlation test to investigate the relationship between overall grade score and AVL vocabulary production. Their results showed no statistically significant correlation overall, only in certain text types, such as comparative analysis essays, argumentative synthesis papers, and response papers. The authors attribute not finding a systematic, linear relationship between the two variables to differences in text types and among the students, as well as variation in the criteria guiding the evaluation of grades (Csomay and Prades, 2018).

Furthermore, In Olsson’s (2015) study, 30 essays were compared in terms of holistic assessment, and differences in AVL production were noted. Results indicates that general academic vocabulary had an impact on assessment. The essays evaluated as better than the example text used for comparison, consisted of a significantly higher number of academic words, than the remaining essays (Olsson, 2015). In her study, the essays judged as weaker, at the same level, and stronger than the text of comparison had an average AVL vocabulary production of 10.57% (SD = 3.37), 11.05% (SD = 3.53), and 13.66% (SD = 3.67), respectively. Thus, this study indicates that there might indeed be a positive relationship between written general academic vocabulary use and assessment of texts.



The relationship between general academic vocabulary knowledge and achieved grades has been investigated in the Norwegian context, but only for receptive knowledge. Skjelde and Coxhead (2020) examined Norwegian L2 English learners in upper secondary. Using the academic sections from two versions of the Vocabulary Levels Test (VLT) by Nation (1983) and Schmitt et al. (2001), the authors used two different mastery levels to determine the participants' meaning-recognition knowledge. One mastery level was 86.7% suggested by Schmitt et al. (2001), while the other level was 96.7%, suggested by Webb et al. (2017). They discovered a moderate association between these scores and course grades:

For participants achieving the 96.7% mastery level, the odds of achieving a higher English grade was over 9 times that of students who did not reach mastery levels, and for those reaching 86.7% mastery the odds of achieving a higher English grade was 4 times that of students who did not reach mastery levels. (Skjelde & Coxhead, 2020, p. 15-16)

In other words, for the participants in their study, receptive knowledge of academic vocabulary was a reliable predictor of English course grades. Examining the relationship between productive general academic vocabulary use and achieved grades among Norwegian L2 learners of English in upper secondary education is therefore a gap in the research, which this study aims to address.

Finally, although academic vocabulary has been emphasized as important for written quality, other factors are likely to impact this, as well. For instance, Dixon et al. (2023) set out to answer empirically how teachers perceive the notion of formality, a concept "integral to academic writing" (p. 161). They found that whether or not writing is perceived as formal is influenced by both linguistic as well as situational features. While academic vocabulary is one example of the former, examples of the latter may be communicative purpose, topic, audience, production circumstances, and so on. In other words, there are aspects influencing the understanding of how formal and academic a text is, besides the use of academic vocabulary.

## 3.0 Methods and Procedures

### 3.1 Research Design

The research design in the current study consists of a combination of quantitative and qualitative approaches, which makes this a mixed methods approach (Johnson & Christensen, 2020). As mentioned in section 2.2, this study takes a corpus-based approach where corpus linguistics is considered a method. Corpora serves as a unique provider of quantitative data (McEnery & Hardie, 2012). Indeed, one might say that by definition, corpus-based methods are quantitative (Gries, 2015). However, when researching writing through corpus studies, Hyland (2010) notes that both qualitative and quantitative methods are used. This is also true for research in corpus linguistics in general, where both analytical techniques are commonly combined (Biber & Reppen, 2015). Mixed methods designs are often defined as either sequential or concurrent (Clark & Ivankova, 2016). In the present study, data were collected in one process. However, data analysis took place in sequences, with subsequent analyses building on the results from the first one.

Mixed methods approaches are increasingly more common in educational research (Brevik & Mathé, 2021). However, perhaps the major challenge associated with this type of research is the requirement of knowledge and skills common to both qualitative and quantitative methods (Creswell & Clark, 2018). Quantitative research skills, in particular, are demanding to familiarize oneself with. In fact, Gries (2015) has noted the lack of “statistical sophistication” (p. 173) as a prominent limitation in corpus studies. Nonetheless, this approach can add precision and provide a more complete understanding of the research topic by combining quantitative and qualitative elements (Ary et al., 2014). The weaknesses associated with one approach can be mitigated by the strengths of the other (Creswell & Clark, 2018). To address the three RQs formulated in this study, different qualitative and quantitative elements were required.

In corpus linguistics, quantitative analytical techniques typically include counting particular linguistic items, such as words (Brezina & McEnery, 2021). For the current study, this was done while answering the first RQ, when counting AVL types and tokens for all texts comprising the corpus. This established the foundation for all further procedures and analyzes in this study. Numeric data was used to describe the extent of general academic vocabulary production in the learner texts, reported in descriptive statistics. Moreover, when examining the overlap of AVL words in the corpus and the assignment briefs and preparation material, similar analytical techniques were used. Further, to answer the second RQ, a qualitative

analysis was included, to determine whether the AVL tokens were used correctly or not. This was possible due to the small size of the corpus, which allowed for a more in-depth analysis. Finally, addressing the third RQ involved using descriptive statistics to evaluate the relationship between the extent of general academic vocabulary use and achieved grades on the English mock exam. Hence, the quantitative components are in the foreground, with the qualitative components integrated. The present study has therefore quantitative priority or weighting (Clark & Ivankova, 2016). Data analyses are described further in section 3.4.

The present study can be described as a type of nonexperimental research. This relates to quantitative research, which may be either experimental or nonexperimental. The former involves a manipulation of one or more (independent) variables and a measurement of the effect this has on another (dependent) variable, while the latter does not (Ary et al., 2014). Yet, in nonexperimental research, one may look at what the relationship is between variables, without manipulation. Notwithstanding the limitations, nonexperimental research can provide invaluable insight in the field of education, where experimental research might not be possible in practical terms (Johnson & Christensen, 2020).

Furthermore, the current study is cross-sectional in terms of the time dimension. In this type of research, data are collected during a short period of time, on several variables (Johnson & Christensen, 2020). Much vocabulary research utilizing corpora of learner language has been cross-sectional, possibly because it allows for considerable quantities of data to be collected in a short time period (Kyle, 2021). While one disadvantage of cross-sectional studies is that the time order is difficult to establish (Johnson & Christensen, 2020), we can safely determine that the learner texts in the current study were written prior to when the grades were given. Moreover, measuring progress in productive use of general academic vocabulary over time was not an objective for this study. Instead, to answer the third RQ, the purpose was to examine a potential relationship between the extent of general academic vocabulary use and achieved grades.

### **3.2 Participants**

To compile the corpus, I reached out to several English teachers and upper secondary schools. This was done through mail correspondence. In the mails sent out, the project was briefly described. I offered to inform classes about the purpose of the project, before getting signed consent by those who wished to participate (see Appendix A). One teacher agreed to this, whereas another teacher distributed consent forms herself and sent me copies. Getting signed

and informed consent, that is, permission to use the texts by those involved, is an ethical consideration, and is an essential part of the corpus creation process (Reppen, 2022).

In total, 32 learners from two different schools gave signed consent to participate in the current study. Two of these learners had not been present on the day of the English mock exam. An additional two learners gave signed consent and had written the mock exam, but the grades they achieved were not available to the teacher, and were thus also excluded. The attrition rate, that is, the decrease in number of participants (Loewen & Plonsky, 2016), was therefore 12.5%. Consequently, 28 participants were included in the present study. In an SLA context, Plonsky (2013) found that the median sample size was 19, in his synthesis of quantitative SLA studies from 1990 to 2010. In other words, the sample size in the present study is relatively small, but this is common in SLA research, even the quantitative studies.

The type of sampling used in this study can be defined as purposive sampling (Ary et al., 2014), purposeful sampling, or criterion-based selection, all of which are used synonymously (Johnson & Christensen, 2020). This approach is commonly used in qualitative research, but may be used in quantitative research as well. Moreover, it is a form of nonrandom sampling, also called nonprobability sampling. Purposive sampling is achieved by recruiting participants according to certain criteria until enough people are obtained (Johnson & Christensen, 2020).

In the context of corpus linguistics, one must therefore define what population to gather samples from (McEnery & Hardie, 2012). The target population in this study, from which the sample was drawn, is Norwegian L2 learners of English in the first year of upper secondary education, in a general studies programme. This is the programme in upper secondary education where learners achieve qualifications for further studies in tertiary education, rather than specific trades. One consequence of using a nonrandom sampling selection method is that not every member of this population had an equal chance of being included in the sample. The use of a nonprobability sampling technique is thus a weak form of sampling compared to probability sampling (Ary et al., 2014). Conclusions from the findings must therefore be drawn carefully, and one must be cautious to make generalizations from the sample to the population. Nonetheless, two advantages of this sampling selection method include “convenience and economy” (Ary et al., 2014, p. 163).

### **3.3 Material**

The English mock exam texts were produced by learners in the target population during the spring of 2023. These texts can be considered high-stakes, given that the mock exam is graded. Further, the texts were collected later the same year, during the fall, along with the grades they achieved on the mock exam. In addition, the preparation material and assignment briefs were collected, which together with the texts and grades, make up the data for the current study. The method of data collection used can be described as the collection of secondary and existing data (Johnson & Christensen, 2020, p. 203). The gathered texts were written for another purpose, at an earlier time, than this study. The grades were also given to the students in the context of their written English mock exams. Thus, all the primary data in this study consist of secondary and existing data. There is a benefit of using written texts that were produced for another purpose than of being used in this study. This benefit is that it can reduce the “Hawthorne Effect” (Landsberger, 1958), which means that participants of a study may change their behavior due to the study taking place. For instance, if learners were told that I wanted to measure their production of academic vocabulary, before writing a text, they might intentionally attempt to produce more of this lexis. This is one challenge associated with vocabulary research (Hirsh, 2010).

#### **3.3.1 Corpus Construction**

##### **3.3.1.1 Balance and Size**

There are several key considerations when building a corpus. First, determining size is important, and this should be guided by two principles: representativeness and practicality (Reppen, 2022). For the current study, these two guiding principles had various implications. First, regarding representativeness, I had to collect enough written mock exam texts for the corpus to sufficiently represent written texts by Norwegian upper secondary L2 learners of English, in the first year of a general studies programme. A representative corpus must consider the principle of balance and the concept of size. Achieving balance entails representing the distribution of language within the population (McEnery & Brookes, 2022). Both balance, and the size needed to accomplish this, is influenced by the RQs and what a given study aims to describe (Durrant et al., 2022). A balanced corpus in the current study could be accomplished by having an equal amount of texts written by males and females, by learners at low, medium, and high proficiency levels, from a variety of socio-economic backgrounds, who only speak Norwegian (L1) and English (L2), to name a few variables which could affect vocabulary usage.

Nonetheless, determining balance and size can be viewed as a “theoretical ideal” (McEnery & Brookes, 2022, p. 38) and “a luxury that we are not always afforded” (p. 39). Yet, certain criteria might be more important than others, such as L1 background, proficiency, and age (Durrant et al., 2022). In the current study, the first criteria used for participant selection was year of instruction, that is, the first year in upper secondary education. This gives a certain indication of the participants’ age, although this was not a selection criteria, nor is information about their age included in the study. It also given certain indications about proficiency level, which is in fact commonly defined by year of instruction (Hulstijn et al., 2010). Equating proficiency level with year of instruction has, however, been criticized (Carlsen, 2012). Moreover, a second criteria was that the participants had to be in a general studies programme, as opposed to a vocational studies programme. Finally, because the texts were sampled from Norwegian L2 learners of English, this gives an indication about L1 being Norwegian and L2 being English.

Practical considerations and limitations will often determine these aspects of the corpus. Therefore, the second guiding principle, practicality, must be considered. For the current study, creating and analyzing the corpus had to be done within the year of writing a master’s thesis, and the collection of texts depended on teachers volunteering to help. Moreover, regarding the size, a very large corpus could be difficult to work with (Durrant et al., 2022). Therefore, the principle of practicality puts certain constraints on the principle of representativeness. A completely representative corpus is often difficult to achieve. Thus, the goal of the corpus creation is instead to gather enough samples for the representation to be accurate (Reppen, 2022).

For the current study, the 28 participants each provided me with a mock exam consisting of three individual texts, resulting in a corpus of 31,109 tokens or running words. The corpus compiled for the current study can therefore be defined as a small specialized corpus. This kind of corpus has an advantage over a large corpus, considering the aims for the study. By purposefully selecting texts, it is likely that the corpus will be more representative of the target population (O’Keeffe et al., 2007). While size should ideally be determined by the RQs, Flowerdew (2004) maintains that pragmatic factors will influence this issue. Moreover, most corpora created by an individual will inevitably be small because of practical limitations (Koester, 2022).

In Table 4 is an overview of the corpus composition in terms of tokens. A complete overview of the composition of both tokens and types in the corpus, is found in Appendix B and C.

**Table 4***Composition of Tokens in the Corpus*

Task	Average length	SD	Total tokens
1	199	24	5,577
2	239	36	6,680
3	673	139	18,852
Total <sup>a</sup>	1,111	143	31,109

<sup>a</sup>Total indicates each mock exam consisting of texts from task 1, 2, and 3.

Table 4 shows that the corpus compiled for this study is significantly smaller than what is typically described as small in corpus linguistics, where the upper limit of small ranges from 250,000 (Flowerdew, 2004) until 5 million tokens (O'Keeffe et al., 2007).<sup>7</sup> In research utilizing a learner corpus, such as the present study, the average size is also typically larger. In their synthesis of learner corpus research from 1991 to 2015, Paquot and Plonsky (2017) found the average corpus size to be around 150 texts and 150,000 tokens, however, with minimum values of 6 texts and 3,286 tokens. The size therefore ranges widely in this type of research.

### 3.3.1.2 Authenticity

An additional criterion to consider when creating a corpus of written texts is authenticity (McEnery & Brookes, 2022). The language should ideally occur naturally, with minimal interference from the researcher. Corpora of learner texts usually have an advantage of having high ecological validity because the language is produced in authentic settings (Durrant et al., 2022). Context is also one aspect to consider which can strengthen the authenticity of the texts. Small specialized corpora have an advantage over larger general corpora regarding the ability to consider contextual information (Koester, 2022). In the current study, collecting the assignment briefs, as well as preparation material for the mock exams, was one way to ensure that context-stripping was avoided.

Each mock exam consisted of four different tasks, and the participants were expected to complete all four. The first task was a digital, multiple-choice task, and was therefore not relevant for the present study. Task number two, three, and four required the learners to write, and all three tasks have been included in the current study. These are referred to as task or text

<sup>7</sup> For spoken corpora sizes of less than a million words is considered small (O'Keeffe et al. (2007, p. 4)

1, 2, and 3. The first task required mediation, the second required interaction, whereas the third required free production in response to a prompt. This was the same for all participants. However, because they were from two different upper secondary schools, the exact task formulations differed. The differences are summarized in Table 5.

**Table 5**

*Overview of Task Differences, with Recommended Word Limit in Parenthesis*

	Mediation task	Interaction task	Free production task
School A	Extract information from an article and write a summary (150-200)	Write a response to a letter in a school newspaper (200-250)	Write a text on the topic “changing values” (500-700)
School B	Extract information from a video and explain it to a student, in writing (150-200)	Write a response to a social media post (150-200)	Write a text on the topic “Knowledge, skills and information in the 21st century” (n.s.)

The inclusion of the different texts aided in broadening the sample diversity, which in learner corpus research tends to be limited, mostly containing argumentative essays (Paquot & Plonsky, 2017). While each learner had written a text in response to task 1, 2, and 3 in a single document, these were saved individually. Yet, the files were named in a way allowing all the texts written by the same learner (i.e. three texts), to be analyzed collectively.

Participants’ names were replaced by a code consisting of two letters, the first indicating which school they belonged to (A or B), and the second to differentiate between the individual participants from the same school (A-Z). Further, each of the three files were given an additional number indicating the task (1-3). Thus, task 1-3 written by the first participant from school A were given the file names *AA1*, *AA2* and *AA3*, respectively. Task 1-3 written by the fourth participant from school B were given the file names *BD1*, *BD2*, and *BD3*, respectively. According to Reppen (2022):

It is always best to create files at the smallest “unit”, since it is easier to combine files in analysis rather than to have to open a file, split it into two texts or more and then resave the files with new names prior to being able to begin any type of analysis. (p. 15)

Creating files in this way allowed for an investigation of the AVL coverage for task 1, 2, and 3, separately and collectively. Paquot and Plonsky (2017) highlight failing to consider results



for particular texts or speakers, and instead only reporting results for corpora in their entirety, as a methodological shortcoming often seen in corpus linguistics. Nonetheless, in certain analyses, all three texts were examined together.

When creating a corpus consisting of learner texts, revision may be necessary for conventional analyses. The corpus texts were received from teachers in .doc and .pdf file formats. Texts received in .pdf file formats were converted into .doc formats. This was done to allow for revision of the texts, such as removal of sources, as well as information indicating task choice. However, when revising the texts, one must keep in mind the principle of authenticity (McEnery & Brookes, 2022). Thus, Reppen (2022) suggests that one should “create an original version, preserving all idiosyncrasies, and a “clean” version that has standardised spelling” (p. 16). Even though I did not standardize spelling in the texts, original versions of the texts were created, as well. Six mock exams collected for this study had been revised by a teacher with notes and suggestions in the texts, which had to be removed when creating the corpus. By mail, I asked what parts of the texts were suggestions made by the teacher, allowing me to edit them back to their original forms.

### **3.3.1.3 Text Selection and Bias**

Finally, Brezina (2018) points out that although texts in a corpus should ideally be sampled non-randomly, “in practice, the selection is guided by text selection principles ... to avoid bias in the selection process” (p. 16). He highlights several biases, some of which were avoided in the present study, and some were not. For instance, the text sample bias involves sampling specific sections of texts in an uneven way leading to an overrepresentation of certain linguistic features found in particular parts of a text, such as the introduction or conclusion (Brezina, 2018). This bias was avoided completely in the current study, by including whole texts in the corpus. Additionally, by including texts from three different tasks, the corpus includes texts written as social media posts and letters in a school paper, and thus, the traditional text type bias was avoided to some extent. This bias is the tendency to only sample texts which typically are sampled in corpus research (Brezina, 2018). However, other biases were not avoided, in particular the topic bias and the self-selection bias. The first bias is especially prominent in small specialized corpora, and arises when a large portion of the texts comprising the corpus is written on the same topic (Brezina, 2018). This has consequences for the type of vocabulary produced in the corpus. The second bias occurs when the gathering of texts is voluntary. Brezina (2018) explains the consequences this might have:

For instance, if we want to create a corpus of classroom writing and ask students to volunteer and contribute their texts, we may end up with texts from highly motivated students that will not reflect the written production of the class as such. (p. 17)

This bias is noteworthy in the present study, and is further reflected in the participants' grades, all belonging to the upper half of the grading scale (see section 3.4.4), suggesting that the self-selection bias was not avoided.

### 3.4 Data Analysis

In the following sections, the data analyses are reported in the order of the RQs. Because measuring the extent of productive general academic vocabulary use (RQ1) established the foundation for all the other analytical procedures, this step is explained in most detail. Despite using the same vocabulary profiling software tool for all three RQs, it is described under the first one. This includes certain challenges and limitations.

#### 3.4.1 Profiling the Corpus for General Academic Vocabulary

To examine the extent of general academic vocabulary production in the written English mock exam texts, the AVL by Gardner and Davies (2014) was used. Academic words were counted using tokens, that is, running words, in order to see the extent of this use. However, it is also interesting to see the degree to which words are repeated. Therefore, counting the number of types, that is, different words, from the AVL, was also done.

Several analytical procedures were carried out to answer the first RQ. First of all, the total instances of AVL items were noted for each of the texts (i.e. 84 individual texts). This provided information on the extent of AVL production for the texts written in response to each of the three tasks in the mock exam separately, and for each of the 28 mock exams in their totality. Additionally, the total number of produced AVL tokens was divided by all tokens making up the corpus, to reveal the overall AVL coverage of the corpus in percentages. The AVL word list was downloaded from [www.academicwords.info](http://www.academicwords.info) as an excel file. Subsequently, I copied the list of words, a total of 3,015, into a *Microsoft Word* document, before converting it to a plain text file (UTF-8 encoded). This way it could be used as a word list ready to be uploaded as a reference list into software for text profiling.

Subsequently, another reference list was created consisting of the first 500 lemmas on the AVL. Similar to the first reference list, this too was converted to a plain text file (UTF-8 encoded) within *Microsoft Word*. The number of AVL types and tokens detected by this list in the corpus of learner texts was noted. These numbers were divided by the all the AVL types

and tokens identified in the first procedure. As a result, we can see what percentage of all AVL items are covered by the first 500 words on the list by Gardner and Davies (2014), alone. The overall AVL coverage of the corpus, provided by this list by itself, was also calculated.

### 3.4.1.1 Software Tool

There are both strengths and weaknesses related to the use of different corpus software. While online tools have an advantage in terms of accessing large corpora, offline tools can be more transparent in that the researcher has a better overview of and control over the data (Anthony, 2022b). When measuring the extent to which general academic vocabulary was used in the written English mock exam texts, I used the software tool *AntWordProfiler* (Anthony, 2023). This is an offline tool, which is free to download. It is the same software tool used by other researchers such as Csomay and Prades (2018), Malmström et al. (2016) and Therovala (2021, 2022, 2023) for purposes similar to the ones in this study (i.e. profiling texts using the AVL). Within this tool, there are three built-in word lists: Coxhead's (2000) AWL and the first and second 1,000 words of the GSL by West (1953). Because there are words within these three lists which also appear on the AVL, they had to be removed before profiling the corpus (L. Anthony, personal communication, december 2, 2023).

However, because the AVL only consists of lemmas and not every member of each lemma, the reference list cannot detect all instances of general academic vocabulary produced in the texts. For instance, words such as *examples* or *tasks* are not picked up on, because only the lemma version of these words, *example* and *task*, are on the AVL. To solve this issue, *TagAnt* (Anthony, 2022a) was used to reduce the words in the corpus texts to their lemma, a process called lemmatization. Therovala (2021, 2022, 2023) used the same software tool in her research, emphasizing that to lemmatize the corpus is necessary when profiling texts using the AVL as it is based on lemmas. In this process, the word *studies* was lemmatized to *study*, and *worst* to *bad*. While this is an easy process (Schmitt, 2010), it does have some limitations, as discussed by Flowerdew (2012, pp. 11-13). When lemmatizing the corpus in *TagAnt*, files were automatically converted into plain text (.txt). Therefore, this is the format which was used when profiling the corpus in *AntWordProfiler*. Figure 1 provides an example of how the software tool was used.

## Figure 1

### Example of Profiling a Text in AntWordProfiler

The screenshot shows the AntWordProfiler interface. On the left, the 'Target Files' section lists 'Test document.txt' with a count of 1. Below it, the 'Reference Lists' section shows 'AVL (Gardner & Davies, 2014)' selected, with a total coverage of 19.92%. Other reference lists include 'ignored' (0.00%) and 'not in lists' (80.08%). The right panel, 'File Profiler', displays the text from 'Test document.txt' (115 types, 261 tokens). The text is color-coded: red for words in the AVL list, green for words not in lists, and black for words ignored. The text includes words like 'furthermore', 'compare', 'two', 'profiled', 'text', 'one', 'which', 'have', 'not', 'be', 'lemmatize', and 'one', 'lemmatized', 'version', 'I', 'notice', 'that', 'certain', 'word', 'be', 'reduce', 'to', 'the', 'wrong', 'lemma', 'for', 'instance', 'know', 'be', 'an', 'avl', 'word', 'whereas', 'the', 'verb', 'be', 'not', 'therefore', 'it', 'go', 'from', 'be', 'detect', 'as', 'belong', 'to', 'the', 'AVL', 'to', 'be', 'exclude', 'from', 'it', 'in', 'fact', 'Martinez', 'and', 'Schmitt', '2015', 'highlight', 'error', 'make', 'from', 'lemmatizing', 'text', 'as', 'something', 'to', 'be', 'cautious', 'of', 'when', 'discuss', 'the', 'use', 'of', 'lemma', 'in', 'word', 'list', 'give', 'that', 'the', 'POS', 'tagging', 'which', 'it', 'be', 'base', 'on', 'can', 'be', 'erroneous', 'to', 'find', 'potential', 'error', 'make', 'in', 'the', 'process', 'of', 'lemmatization', 'I', 'track', 'which', 'avl', 'word', 'be', 'lose', 'by', 'profile', 'the', 'corpus', 'before', 'and', 'after', 'lemmatization', 'and', 'observe', 'the', 'difference', 'in', 'avl', 'word', 'these', 'include', 'word', 'such', 'as', 'know', 'educate', 'datum', 'all', 'of', 'these', 'word', '34', 'in', 'total', 'be', 'manually', 'search', 'for', 'in', 'the', 'lemmatized', 'version', 'of', 'the', 'corpus', 'and', 'change', 'be', 'make', 'to', 'include', 'they', 'however', 'they', 'be', 'also', 'check', 'for', 'correct', 'pos', 'to', 'rectify', 'the', 'issue', 'of', 'overcount', 'indeed', 'most', 'of', 'the', 'other', 'word', 'exclude', 'after', 'lemmatize', 'the', 'corpus', 'be', 'not', 'erroneous', 'and', 'thus', 'rightfully', 'exclude', 'such', 'instance', 'include', 'several', 'verb', 'which', 'be', 'equivalent', 'to', 'adjective', 'include', 'on', 'the', 'avl', 'when', 'appear', 'in', 'a', 'particular', 'inflect', 'form', 'an', 'example', 'be', 'receive', 'an', 'avl', 'item', 'when', 'use', 'as', 'an', 'adjective', 'properly', 'reduce', 'to', 'the', 'lemma', 'receive', 'and', 'therefore', 'correctly', 'exclude', 'from', 'the', 'avl', 'coverage'.

*Note.* The words color-coded in red belongs to the uploaded reference list, the AVL.

Further, there are several aspects to consider when analyzing a corpus. These include deciding whether or not numbers and punctuations should be counted as words, and whether hyphenated words and contractions should be counted as one or two words (Durrant et al., 2022). Determining these issues depend on how the software tool counts words. In *AntWordProfiler*, I tried to profile a few documents to establish how words are counted by default. The example *can't* consists of two tokens and two types, which are *can* + *t*. In other words, contractions are counted as two words, but the software tool only counts what is after the apostrophe, not the actual word it represents. This could skew the counting of tokens, because every contraction would be counted as two tokens. To cope with this, one option would be to expand every contracted word, so that *he's* is written as *he is*. However, this would skew the overall token count similar to the default way of counting contractions. Thus, I changed the token definition within the software tool itself, so that contracted words are only counted as one token, and one type. However, because contracted words contain

different types of apostrophes, such as ` or ‘ or ’, this caused some issues, because only contractions with a particular type, namely ` as in *he`s*, was detected. To solve this, I changed all apostrophes in the corpus texts to the one which the software tool managed to detect without problems, so that *he’s* was changed to *he`s*.

Moreover, in *AntWordProfiler*, punctuations and numbers are not counted, whereas one can choose whether or not to count capitalized words as separate words. I decided not to count capitalized and lowercase words separately, as making such a distinction would not add any important information in this study. Additionally, hyphenated words such as *real-life* are counted as two tokens and two types, that is, *real* + *life*, by default. This could potentially have an impact on the AVL coverage as the word list contains 99 hyphenated words (see Appendix D for a full list). Because the token definitions of the AVL and *AntWordProfiler* are different by default, editing how tokens are counted in *AntWordProfiler* was necessary so that hyphenated words are included (L. Anthony, personal communication, december 2, 2023). The way this was done was by replacing the original token definition “[p{L}]+” with “[p{L}`-]+” so that hyphenated words, such as *real-life*, are counted as one token, and the same with contractions.

However, only one of the hyphenated words on the AVL, *long-term*, is found among the 500 most frequent words. As discussed in section 2.6.1, this is the portion of the list which makes up the majority of AVL words used in both the academic section of the COCA, and in university writing (Malmström et al., 2016). The likelihood of learners using hyphenated words found on the AVL, was therefore considered to be low. Nevertheless, editing the default token definition ensured that any instances of hyphenated AVL words would be detected. One final aspect to point out is that after lemmatizing the texts in *TagAnt*, hyphenated words were changed from *real-life* to *real - life*. In other words, an additional space was added on each side of the hyphen. To cope with this, I manually edited each lemmatized plain text version of all the corpus texts so that the additional spaces between the hyphens were removed, in order to allow for the words to be detected in *AntWordProfiler* as individual words.

### 3.4.1.2 Challenges

One significant limitation is associated with the use of an untagged corpus, which is that a distinction is not made between POS. This limitation is addressed by Malmström et al. (2016) who write that an untagged corpus “does not distinguish between words like study, n., which

is on the AVL, and study, v., which is not” (p. 55). The initial profiling of the texts revealed that the words *class* and *mean* were among the most common AVL words in the corpus. However, *class* was never used as a verb, and *mean* was seldom used as a noun, the POS each word has on the AVL. The use of an untagged corpus therefore leads to an overcounting of AVL words. This is a challenge associated with the use of lemmas as the counting unit, given that “current vocabulary profiling software does not distinguish orthographically identical word forms” (Stoeckel et al., 2020, p. 605). In order to cope with this issue, I manually searched for all AVL items detected in the initial profiling of the corpus, and correct POS was determined for all individual occurrences. In the cases where words from the wrong POS were included as AVL words, these were simply removed from the AVL count. After excluding words that were wrongfully included, such as *use* (v) or *view* (n), the overall AVL items in the corpus was reduced by 598 tokens. Given that these instances were manually checked and evaluated for correct POS by the author of the present study, errors may have occurred in this process, such as excluding a correct AVL word or including an incorrect one. This is an issue of reliability, further discussed section 3.5.

Furthermore, after comparing two profiled texts, one which had not been lemmatized and one lemmatized version, I noticed that certain words were reduced to the wrong lemma. For instance, *known* (j) was reduced to *know* (v). In the corpus texts, *known* was often used as an adjective, which is an AVL word, whereas the verb is not. Therefore, it went from being detected as belonging to the AVL, to being excluded from it. Indeed, the POS tagging that lemmatization is based on, can be erroneous (Brezina, 2018; Martinez & Schmitt, 2015). Moreover, Paquot and Plonsky (2017) underline that learner language can impede successful POS tagging and lemmatization.

To find potential errors made in the process of lemmatization, I traced which AVL words were lost by profiling the corpus before and after lemmatizing and observed the difference in AVL words. These included words such as *known* (j), *educated* (j), and *data* (n). All of these words were manually searched for in the lemmatized version of the corpus, and changes were made to include them. However, they were also checked for correct POS to rectify the issue of overcounting. Indeed, most of the words excluded after lemmatizing the corpus were not erroneous, and thus rightfully excluded. Such instances include several verbs, which were equivalent to adjectives included on the AVL when appearing in a particular inflected form. An example is *received* (j), an AVL item when used as an adjective, properly reduced to the lemma *receive* (v), and therefore correctly excluded from the AVL coverage. This process led to the inclusion of an additional 19 AVL vocabulary tokens.

Thus, the inclusion of non-AVL items as a result of using an untagged corpus can be more problematic than the exclusion of AVL words erroneously omitted when lemmatizing the texts. Nonetheless, both issues were able to be coped with due to the small corpus size. However, another limitation appeared along the same lines as the challenges discussed above. The fact that none of the words were tagged with POS introduced the difficulty of distinguishing between AVL words of identical form but different POS. Several words on the AVL occur more than once, for instance as both a verb and a noun. To exemplify this, consider the word *result*. This word occurs on the word list as both a verb and a noun. Such words are grouped together in the profiling of AVL words used in the corpus texts. As a result, when the learners have produced the word *group* in their texts, for instance, it can be difficult to assert whether this is the noun *group*, a very common AVL word listed as number 2, or the verb *group*, a less common AVL word listed as number 1,339.

However, the objective in the current study was not primarily to establish which particular academic words are used by the learners, but to what extent they rely on such words. Because this information was extracted by counting AVL tokens and dividing this with the total number of tokens, the AVL coverage was unaffected by which type of AVL items were used. Nonetheless, this limitation is worth underlining, and to keep in mind whenever specific words are highlighted. A total of 100 words on the AVL appear in more than one word class, which are listed in Appendix E. Half of these are produced in the corpus in the present study, listed in Appendix F. Thus, to some extent, the list of AVL types produced in my corpus, shown in Appendix G, consists not only of lemmas, but also flemmas. However, words are only grouped together if the POS for each word is on the AVL.

### **3.4.2 Examining the Impact of Assignment Briefs and Preparation Material**

Furthermore, RQ1a aimed to examine a potential link between the AVL vocabulary produced in the corpus, and the vocabulary found in the assignment briefs, as well as the preparation material. The AVL types detected in the first procedure (see Appendix G) were made into a reference list. This was converted to a plain text file (UTF-8 encoded) in *Microsoft Word*. Subsequently, it was uploaded to *AntWordProfiler* and used to profile two files. The first file consisted of the assignment briefs from both schools, containing the prompts, as well as texts the learners had to read in the process of writing their mock exam. This included one video that some of the learners had to watch in the mediation task, which I transcribed. I will refer to this as part of the assignment brief. Thus, whenever the assignment brief is referred to, it is important to keep in mind that this includes prompts and source texts, as well as one video.

The second file contained the same as the first, but also the preparation material, from both schools. This consisted of texts on the overall topic of each mock exam.

Both documents were lemmatized in *TagAnt*, and the removal of additional hyphens were done to ensure that hyphenated words were detected. After profiling the files with the AVL types produced in the corpus, correct POS was evaluated to avoid overcounting. This was a significantly easier process than profiling the entire corpus of learner texts, because in this procedure, it was enough to detect one instance of the AVL type in the correct POS, for it to be included in the analysis. In addition, the AVL types that were detected as having been erroneously excluded because of errors in the lemmatization process when profiling the corpus (see section 3.4.1.2), were manually searched for in the unlemmatized versions of the documents profiled in this procedure. Three tokens were included as a result. The word count in *Microsoft Word* revealed that the file containing the assignment briefs consisted of 3,823 words, whereas the one containing the preparation material in addition to the assignment briefs, consisted of 9,849 words.<sup>8</sup>

While it would have been interesting to examine how many of the 3,015 lemmas on the entire AVL by Gardner and Davies (2014) were present in these files, it was limited to the words produced in the corpus, in order to detect any overlap. The intention was to investigate how many of the AVL items produced in the corpus, were also present in the texts the learners had to read, either to write their mock exam, or in preparation for it. Therefore, the amount of AVL tokens in the corpus of learner texts, achieved by the AVL types found in the two files profiled in this procedure, was examined.

This was done by creating two additional reference lists, each containing the AVL types detected when profiling the two files (assignment briefs and assignment briefs + preparation material). These were converted to plain text files (UTF-8 encoded) in *Microsoft Word*, uploaded to *AntWordProfiler* and used to profile one file containing all the AVL tokens detected in the entire corpus of learner texts.

It should be noted that in this analysis, the assignment briefs and preparation material provided to the participants from both schools, were combined. This was considered reasonable because of the unequal number of participants from each school. However, alternative steps in this procedure might have provided additional insights. For instance, by

---

<sup>8</sup> The process of ensuring that contracted words are counted as one instead of two tokens, similar to the procedure done when profiling the corpus of learner texts, was not carried out when profiling these documents. This would have been time-consuming without serving any purpose. Because of this, the number of words was reported instead of tokens, as the latter would have been inaccurate (too high).



creating two sub-corpora, one containing the learner texts from school A and one containing the learner texts from school B, examining whether or not learners from each school relied on the vocabulary from the assignment briefs and preparation material to different extents, would be possible. Moreover, one could create multiple reference lists, distinguishing between the different parts of the assignment briefs. However, a similar procedure including the preparation material would have been difficult, as this contained texts on the overall topic of the mock exam, not for each separate task. Nonetheless, the decision to combine the assignment briefs from both schools, in their entirety, was done to provide a general overview of the potential influence of the general academic vocabulary found here, on the vocabulary produced by the learners. Similarly, including the preparation material from both schools, and combining this with the assignment briefs, provided information regarding how much additional general academic vocabulary this contributed with, on the whole.

### 3.4.3 Evaluating Correct Use

To answer the second RQ, I relied on Hasselgren's (1994) taxonomy to evaluate whether the general academic words produced by the learners were used correctly or not. I took a lenient approach where I only utilized the category of semantic dissonance (see section 2.3.2). This involved a qualitative analysis where the words belonging to the AVL were examined in their contextual environment. The objective was to determine whether or not the academic words produced were used correctly in terms of their meaning within the given context. As Mahan (2013) points out in her study, "[t]he only way to pinpoint semantic dissonance in free production texts is through context" (p. 39). The software tool used for profiling, *AntWordProfiler*, made this easy by color-coding the output, showing where the AVL words occurred in the texts (see Figure 1). Therefore, this tool was used when carrying out this analysis, as well.

When reporting on correct use of AVL words, this is presented in terms of a percentage out of all the AVL words produced by each learner, in all three tasks. In other words, if a learner produced 100 words from the AVL, and two words were used incorrectly, the number reported under correct usage would be 98%. While it might appear obvious which words are semantically incorrect, certain instances were difficult to establish. Words could be semantically correct, and simultaneously be incorrect in terms of other aspects, such as having the wrong inflection (e.g. plural/singular). Moreover, misspelled words were completely excluded from the AVL count and were therefore not evaluated for correctness. This is

because of how the profiling tool works, where words need to be orthographically similar to the words on a given reference list, to be detected as belonging to that list.

All the general academic vocabulary produced in the corpus was evaluated by the author of the present study. However, the ones I determined were incorrect were subsequently evaluated by two other faculty members from the linguistics department at the university in Bergen. This process strengthens inter-rater reliability (see section 3.5). Ideally, as Loewen and Plonsky (2016) underline, all data should be double-coded, however, this is unusual. The common practice is instead to have a small portion of the data double-coded. Common indices used to estimate agreement include Cohen's Kappa or percent agreement (Loewen & Plonsky, 2016). Given the small number of data that were double-coded, the latter index was used. This is a statistic which can be directly interpreted and is easy to calculate (McHugh, 2012).

Considering that the corpus compiled for the present study comprises 31,109 tokens, and only the general academic vocabulary tokens within this corpus were evaluated for correct use, the number of semantically dissonant instances was expected to be low. Nonetheless, it was interesting to investigate whether or not the general academic vocabulary produced was used correctly or not. Moreover, this was one way to strengthen the validity of the findings in the present study. Unlike the words that were instances of overcounting, due to being another word class than the lemma on the AVL (see section 3.4.1.2), the general academic vocabulary evaluated as incorrect is included when the AVL vocabulary coverage of the corpus is reported.

#### **3.4.4 Investigating the Relationship with Achieved Grades**

To answer RQ3, the relationship between the extent of AVL vocabulary use and achieved grades was examined. Statistical measures, such as significance testing, is commonly used in corpus linguistics, to get a better understanding of information derived from a corpus (McEnery & Hardie, 2012). Initially, the intention was to rely on inferential statistics, and significance testing, in particular. However, due to the small sample size selected non-randomly, and failing to avoid the self-selection bias in particular, resulting in only high grades, the use of inferential statistics was limited. When the sample size is small, it might be more suitable to rely on descriptive statistics (Plonsky, 2014), which is what was used to assess the relationship between AVL production and achieved grades.

In order to answer the third RQ, the grades had to be grouped together to allow for further analyses. The grading system in Norwegian upper secondary education utilizes the

numbers 1 to 6 where 6 is the highest grade. However, the grades assigned to the learners on the mock exams collected for this study also consist of instances such as 4+ and 5-, as well as 5/6 and 5/4. The grades among the participants in this study are high, ranging from 4- to 6. To my knowledge, an established way to convert these grades into numerical equivalents does not exist. In order to cope with this issue, different grades were merged together into two groups: middle (mid) and high grades. The mid group ( $n=15$ ) included grades 4-, 4, 4+, 4/5, 5/4, and 5-, whereas the high group ( $n=13$ ) included grades 5, 5+, 5/6, 6/5, and 6. While the division might seem arbitrary, it nonetheless grouped together grades at approximately the same interval, that is, from 4- to 5- in the mid group, and from 5 to 6 in the high group. At the same time, it resulted in nearly the same number of participants in each group. The corpus of learner texts was divided into two sub-corpora based on these grades, to answer this RQ.

The extent of general academic vocabulary use in the written mock exam texts was the independent variable, whereas achieved grade was the dependent variable. When assessing the relationship between AVL production and achieved grades, I predicted a positive relationship. In other words, that learners who produced texts with higher AVL coverage, also achieved higher grades. SPSS version 29 was used to create a boxplot and scatterplot to visualize a potential relationship.

Furthermore, other analytical procedures were carried out, similar to those used when answering RQ1 (see section 3.4.1) and RQ1a (see section 3.4.2). First, the top 500 words on the AVL were utilized to profile the texts written by each group to explore any differences in the use of these words, based on grade. Second, the AVL types found in the assignment briefs alone and then in combination with the preparation material, were used as reference lists to profile the texts written by each group. The intention was to investigate any differences in the reliance on these words.

### **3.5 Validity and Reliability**

Validity and reliability are two concepts related to study quality which need to be viewed in relation to each other. In general terms, validity refers to whether we measure what we intend or claim to measure, while reliability refers to the level of consistency (Ary et al., 2014). In the current study, a word list was used to profile the corpus of learner texts. Word lists need to be evaluated for how valid they are for particular purposes, and how reliable or consistent they are, as well (Schmitt & Schmitt, 2020). The validity of a word list can be evaluated based on claims about its usefulness, as well as the corpus or corpora it is derived from. Regarding

the latter, both the material and size are important aspects to consider. The usefulness of the AVL, according to Gardner and Davies (2014), is in representing a core academic vocabulary. As discussed in section 2.5.2, the corpus that the AVL is based on is significantly larger and contains texts from more disciplines, compared to the corpus that the AWL is based on.

Further, corpus size also impacts the reliability of the word list utilized. The question is whether the corpus that the word list developers rely on is large enough to capture the words they intend to capture (Schmitt & Schmitt, 2020). In simple terms, if increasing the corpus size does not lead to the inclusion of additional items on the word list, then it is reliable and consistent. Two other aspects influencing reliability of a word list are the age of the corpus, and the degree of specificity in regard to the word list (Schmitt & Schmitt, 2020). Age has to do with whether the source corpus is sufficiently updated to include current words. Again, the AVL is not only a more recent word list, but given that the AWL was built on top of the GSL, the AVL is a considerably more updated word list overall. Further, the specificity of the word list must be considered for the purposes of its use. In the current study, the purpose is to profile written texts for the productive use of general academic vocabulary. Thus, a list such as the AVL is at an appropriate level of specificity for the purposes of this study.

Moreover, confirming that academic words are in fact detected is essential when discussing issues of both validity and reliability. Before profiling the corpus texts in the present study, I tried profiling a document containing all 3,015 AVL words to see if it showed 100% coverage. This was done to establish that the reference list of AVL words which I used could identify these words accurately and consistently. This process included changing the token definition in *AntWordProfiler* to ensure that hyphenated and contracted words were counted as individual tokens (see section 3.4.1.1). Furthermore, lemmatizing the texts using *TagAnt* was important in order to identify all the AVL items used in the texts. I also attempted to identify AVL vocabulary that was erroneously lemmatized to the wrong lemma (see section 3.4.1.2).

In addition, evaluating correct use of the AVL words produced in the corpus was one way to strengthen validity. As Csomay (2020) suggested, this may provide a more accurate analysis of general academic vocabulary production. However, this very process intended to strengthen validity might itself cause certain issues. Errors can be made in the process of manual coding, which, according to Larsson et al. (2020), may be either systematic or random:

Manual coding might be compromised for a variety of reasons, both systematic (e.g. due to ambiguity of the coding scheme, inadequate coder expertise or training and/or coder bias) and random (e.g. due to coder fatigue and/or typing errors); all of these will introduce inaccuracies into our results. (p. 238)

The authors therefore emphasize the importance of inter-rater reliability, in particular, especially for studies in LCR where manual coding is common (Larsson et al., 2020). This procedure examines the level of agreement or consistency between different raters, when they analyze the same data base on certain criteria (Larsson et al., 2020).

In the current study, inter-rater reliability was strengthened during the process of determining correct use of the AVL words, to some extent (see section 3.4.3). When profiling the corpus for AVL production (RQ1), manually evaluating all AVL words initially detected for correct POS to avoid overcounting, was a less subjective process than evaluating correct use. Nonetheless, this analytical procedure might have been more likely to be affected by random errors (Larsson et al., 2020), considering the repetitive process of 1) determining the word class of an AVL word, 2) identifying the word on the AVL itself, and 3) deciding if these are the same, before 4) manually removing the word if it was not the same word class. However, an inter-rater reliability procedure involving double-coding of a small portion of this data would likely have shown a high degree of agreement, as it is unlikely other raters would evaluate POS differently. Thus, such a procedure might be misleading, given that coding errors were more likely to occur due to random mistakes such as fatigue, than for systematic reasons.

Another issue affecting validity in particular, is whether or not the tasks on the mock exam actually elicit or ask for the use of an academic language. As discussed in section 3.3.1.2, the mock exam consisted of four different tasks. The first was irrelevant to look at, as it was a multiple-choice task. However, the three remaining tasks all elicited writing and were therefore included as data. Yet, it is important to be transparent about the fact that these tasks might elicit an academic language from the learners to various extents. I would consider it apparent that task 3 which is a free production task requires an academic language to a greater extent than task 1 and 2, which are mediation and interaction tasks, respectively. Moreover, there may be differences between the schools. Even though the tasks are similar, the exact details differ. For instance, in the interaction task in school A, the learners had to respond to a letter in a school newspaper. This letter appears more formal than the social media post in the equivalent task in school B. Finally, tasks given in school A specify certain requirements. In task 1 it states that the learners should “[w]rite a coherent text”. In task 2 it states that they

should “strive to use proper language”. No such expectations are made explicit in the assignment brief provided at school B.

Finally, describing the steps taken in the analytical procedures, as well as explaining which software tools are used, are important to strengthen reliability. There is an unfortunate tendency to not report what programs are used to analyze corpora containing learner language. As many as 34% of the studies investigated in Paquot and Plonsky’s (2017) synthesis of quantitative studies in LCR left out this information, which according to them “should be reported for the sake of transparency and replicability” (p. 70). In addition, I would argue that explaining the challenges which occur and how they are dealt with, contributes to this transparency. In the current study, detailed descriptions about the steps taken when compiling and analyzing the corpus strengthen transparency. This allows for reliable replication of the study.

### **3.6 Ethics**

Before recruiting participants in the process of collecting data, the project was registered in system for risk and compliance (RETTE) during the fall of 2023. This is the University of Bergen’s system to monitor responsible processing of personal data in research projects. Thus, registration was done in consideration of personal data gathered through the data collection. Although participants would remain anonymous, and only their written mock exam texts and grades were gathered, it was necessary to register the project in RETTE and get signed consent from participants. Further, in order to ensure anonymity, wherever the names of the students or their school appeared in the corpus, these names were replaced. In addition, file names were replaced by codes (see section 3.3.1.2) used consistently throughout all analytical procedures.

## 4.0 Results

Results are presented in the order of the RQs. First, the extent to which the learners used general academic vocabulary from the AVL in their written production, is reported. This includes the proportion of AVL vocabulary items in the corpus as a whole, the average for each mock exam, as well as the average for the three separate tasks comprising each mock exam. It also includes the proportion of AVL vocabulary made up by the first 500 words on the AVL word list. Second, the overlap of AVL vocabulary in the corpus and the assignment briefs and preparation material, is reported. Third, results from the analytical procedure of evaluating correct use of the AVL vocabulary, are presented. Finally, the relationship between AVL vocabulary production and achieved grades is presented through descriptive statistics and graphic visualizations of this data.

### 4.1 General Academic Vocabulary Production in the Corpus

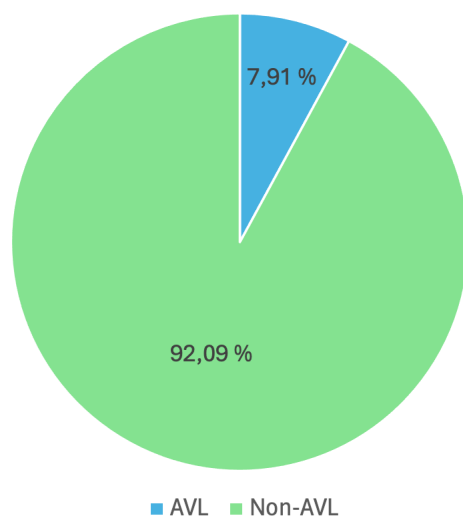
The results from profiling the corpus in *AntWordProfiler* reveal that vocabulary from Gardner and Davies' (2014) AVL make up 7.91% of tokens in the entire corpus. AVL types make up 16.65% of all word types produced in the corpus. This is presented in Table 6.

**Table 6**

*Overall Corpus Composition and AVL Coverage*

	Corpus tokens	Corpus types	AVL tokens	AVL types
Raw numbers	31,109	3,033	2,461	505
Percentages	100	100	7.91	16.65

Furthermore, these results are illustrated in Figure 2. This figure shows the overall coverage of general academic vocabulary from the AVL in the corpus, compared to vocabulary that may be either general or technical, which is simply defined as non-AVL words.

**Figure 2***AVL Token Coverage of the Corpus*

Furthermore, each participant produced an average of 7.96% (SD = 3.33) general academic words from the AVL in their entire mock exams. This proportion of AVL vocabulary differs between the three tasks. For a detailed overview of the production of AVL tokens by all the participants, see Appendix H (percentages), as well as Appendix I (raw numbers). AVL types produced by the learners in each task is shown in Appendix J (raw numbers).

Descriptive statistics of AVL use based on task, both in terms of percentages as well as raw numbers, are presented in Table 7 and 8.

**Table 7***Distribution of AVL Tokens in the Corpus in Percentages*

	Mean	SD	Min	Max	Range
Task 1	8.07	3.24	2.35	15.05	12.70
Task 2	5.76	2.52	1.51	12.44	10.93
Task 3	8.76	4.95	0.43	21.95	21.51
Total	7.96	3.33	2.31	15.92	13.61



**Table 8***Distribution of AVL Tokens in the Corpus in Raw Numbers*

	Mean	SD	Min	Max	Range
Task 1	15.89	6.46	5	31	26
Task 2	13.64	5.82	3	25	22
Task 3	58.36	35.64	3	176	173
Total	87.89	37.68	26	196	170

Thus, when analyzed separately, the average AVL production is 8.07% (SD = 3.24), 5.76% (SD = 2.52), and 8.76% (SD = 4.95) for task 1, 2, and 3, respectively. The proportion of AVL tokens is highest in the texts for task 3, and lowest in the texts for task 2. However, the results indicate that all the tasks elicited general academic vocabulary to some extents. Whether or not the tasks asked for or elicited an academic language was highlighted as an issue of validity (see section 3.5). Further, AVL use is most widely dispersed in the texts for task 3, where the standard deviation is highest. Interestingly, in these texts, the range in AVL use is 21.51, which is approximately twice of the range in the texts written for task 1 or 2.

Moreover, as shown in Table 8, most AVL words are produced in task 3, on average. This distribution of AVL words across tasks is further illustrated in Table 9.

**Table 9***AVL Coverage for Each Task Making up the Mock Exams*

	Task 1	Task 2	Task 3	Total
AVL raw numbers	445	382	1,634	2,461
% of all tokens <sup>a</sup>	7.98	5.72	8.67	7.91
% of AVL tokens <sup>b</sup>	18.08	15.52	66.40	100

<sup>a</sup>Percentages of all tokens in the corpus ( $n=31,109$ ). <sup>b</sup>Percentages of all AVL tokens in the corpus ( $n=2,461$ ).

This table illustrates the fact that the majority of AVL items produced in the corpus occur in the texts written for task 3 in the mock exam. In fact, two-thirds of the AVL tokens (66.40%) are found here. This is not surprising, given that these texts have an average length of 673 (SD = 139) tokens compared to 199 (SD = 24) and 239 (SD = 36) for the texts written in response to task 1 and 2, respectively (see Table 4).

Further, the corpus was profiled using only the first 500 lemmas on the AVL. This analytical procedure revealed a total coverage of 2,025 AVL tokens. This makes up 82.28% of all the AVL tokens produced in the corpus, which by itself provides a coverage of 6.51%

of all running words in the corpus. In terms of AVL types, the first 500 words on the AVL make up 286 out of the 505 AVL types produced in the corpus, which is 56.63% (see Appendix K). In other words, a considerable amount of general academic vocabulary is achieved by this small portion of the AVL.

#### 4.2 Impact of Assignment Briefs and Preparation Material

Results from the analytical procedure of comparing the assignment briefs and preparation material to the AVL words produced in the corpus of learner texts, show extensive overlap. In the profiling of the document containing the assignment briefs, 149 of the 505 AVL types produced by the learners were detected (see Appendix L). These word types account for 1,556 of the AVL tokens used by the learners, which makes up 63.23% of the total 2,461 AVL tokens in the corpus. In Table 10 we can see the portion of AVL tokens produced in each task which is made up of types from the assignment briefs.

**Table 10**

*AVL Tokens Achieved by AVL Types from the Assignment Briefs*

	AVL tokens in raw numbers	% of AVL tokens in each task
Task 1	315	70.79
Task 2	218	57.07
Task 3	1,023	62.61
Total	1,556	63.23

Table 10 illustrates that the AVL types in the assignment briefs make up different portions of the total number of AVL tokens, in the texts produced in the various tasks. The greatest overlap is seen in task 1, where 70.79% of the AVL tokens are achieved by types from the assignment briefs. In task 2, the overlap is the smallest, where 57.07% of the AVL tokens are made up of types from the assignment briefs.

Further, profiling the document which contained the preparation material, in addition to the assignment briefs, revealed an even greater overlap. A total of 243 of the 505 AVL types produced by the learners were detected (see Appendix M), which account for 1,893 of all the AVL tokens. This makes up 76.92% of the total 2,461 AVL tokens in the corpus. The portion of AVL tokens in each task, made up of types from the assignment briefs and preparation material, is shown in Table 11

**Table 11**

*AVL Tokens Achieved by AVL Types from the Assignment Briefs and Preparation Material Combined*

	AVL tokens in raw numbers	% of AVL tokens in each task
Task 1	359	80.67
Task 2	265	69.37
Task 3	1,269	77.66
Total	1,893	76.92

Thus, the number of AVL tokens achieved by types in the assignment briefs and preparation material, increases in the texts written for all three tasks. Again, we see that these types make up the greatest portion of AVL tokens in task 1, and lowest in task 2. Nonetheless, the majority of the general academic vocabulary tokens stem from types found in the assignment briefs and preparation material, irrespective of which task.

The influence of the assignment briefs, as well as the preparation material, on the productive use of general academic vocabulary, is illustrated by this overlap. The large number of tokens these types cover in the corpus, makes this influence even more evident. These findings also point to the fact that these AVL words, found in either the assignment briefs alone, or in combination with the preparation material, are among the AVL types produced most frequently by the learners. In fact, 39 of the 50 most frequent AVL types in the corpus (see Appendix G) appear in the assignment briefs. Furthermore, 44 of the 50 most frequent AVL types in the corpus appear in the assignment briefs and preparation material combined.

Along these lines, if we look at the AVL words produced most frequently, we discover another interesting finding. While some AVL types are produced several times, other types are produced less. On the one hand, the top 50 AVL types from Appendix G cover 1,256 AVL tokens which make up 51.04% of the total amount of AVL tokens. Put differently, approximately 10% of the AVL types used by the learners cover approximately 50% of the AVL tokens in the entire corpus. This illustrates the skewness in how often the general academic words are produced. In fact, on the whole, most AVL types only occur a few times in the entire corpus. 206 AVL types occur once, 76 AVL types occur twice, and 60 AVL types occur three times. In other words, 342 of the 505 AVL types produced in the corpus are used three times or less. Together, these types make up 21.86% of all AVL tokens.

### 4.3 Correct Use of General Academic Vocabulary

The 2,461 AVL tokens produced in the corpus were evaluated for correct use, with the category of semantic dissonance in Hasselgren's (1994) taxonomy as the guiding principle determining incorrect use. The words were evaluated in their contextual environment, where wrongness in meaning was the only criterion. Based on the words preceding and succeeding each general academic word, the intended meaning was assumed. Results revealed a modest amount of productive general academic vocabulary used incorrectly. In total, merely 20 AVL tokens were evaluated as incorrectly used, due to the effect of semantic dissonance. Further, these tokens were produced by 15 of the participants. In other words, approximately half of the participants used none of the general academic words incorrectly.

On average, 99.25% of the AVL words produced in the corpus were used correctly, with a standard deviation of 0.79. The descriptive statistics are summarized in Table 12.

**Table 12**

*Correct Use of AVL Words Produced in the Corpus in Percentages*

	Mean	SD	Min	Max	Range
Task 1-3	99.25	0.79	97.92	100	2.08

For certain AVL words, it was possible to determine the intended meaning based on the task which the learners were responding to. For instance, one example from the corpus of an AVL word deemed incorrect, is *achieves* in the following sentence:

- (1) This is proven by looking at the old **achieves**.

The intended meaning was interpreted to be the noun *archives*, as this word occurred in the text the learner was summarizing on the mediation task. Other examples were more difficult to determine, given that they could be incorrect along other lines than wrongness of meaning. In the following sentence, the AVL word *apply* was not evaluated as incorrect, even though the correct inflection would be the singular form *applies*:

- (2) Teens are always checking their phone, but this utterance also **apply** for the grown-ups.

Furthermore, it is worth noticing that the AVL word *critique*, produced twice in the entire corpus, was considered semantically incorrect in both contexts where it occurred. It was used in the following constructions:

- (3) [T]hey [FIFA] get so much **critique**.
- (4) [T]he younger generation often receive **critique** about [their spending].

In both of these instances, the intended meaning was interpreted to be *criticism*. All the words evaluated as incorrect are shown in Appendix N. The following sentences highlight a selection of examples:

- (5) The decision whether it is good to change the past is complex as it **effects** people from all sides.
- (6) The values of a person will have great impact on the persons decisions, and therefore **effect** one's life.
- (7) Lastly, one should also make sure one does not recycle **previous** submitted works, as this is seen as self-plagiarism.
- (8) Values are often passed down through family, **development** from own experiences as well as history.
- (9) That is why it is super important to be well educated about these kinds of things, so that we are able to **provide** and avoid them from happening again.

In example 5-9, the intended meaning was interpreted to be *affects*, *affect*, *previously*, *developed*, and *prevent*, in the respective sentences. The produced and intended words in example 7 (*previous/previously*) might seem similar to the ones in example 2 (*apply/applies*), which was not considered incorrect. Nonetheless, *previously* is an adverb and *previous* is an adjective, whereas *apply* and *applies* are both verbs. Thus, the former differ in word class, whereas the latter differ in inflection only. Still, one could argue that the meaning in both examples is only affected marginally, and that none of the words should be considered incorrect based on the principle of semantic dissonance. Nonetheless, *previous* was evaluated as incorrect while *apply* was not. As this illustrates, not all instances were easy to evaluate.

The procedure intended to strengthen inter-rater reliability (see section 3.4.3), showed a high level of agreement. In some cases, the two other linguistics faculty members evaluated a word as uncertain or not entirely incorrect. When calculating the percent agreement

(McHugh, 2012), such cases, and the ones evaluated as correct, were scored 0, and the words rated as incorrect were scored 1. The inter-rater reliability score was 0.85, or 85%. The agreement on incorrectness in terms of semantic dissonance ranged between 66% and 100% for every AVL word evaluated. Thus, for all 20 words, two thirds, that is, me and one of two external raters, agreed that the AVL word was incorrect. Consequently, no changes were made to the words evaluated as incorrect. This process nonetheless elucidates the important role the researcher has on the analysis of corpus data, given that others may interpret them differently.

#### 4.4 Relationship with Achieved Grades

The aim of the third RQ was to investigate any potential link between the use of general academic vocabulary and achieved grades. The participants were grouped together into two categories, differentiating between mid and high grades. AVL use is not reported for the three individual texts separately but combined for each of the 28 learners. Descriptive statistics are presented in Table 13.

**Table 13**

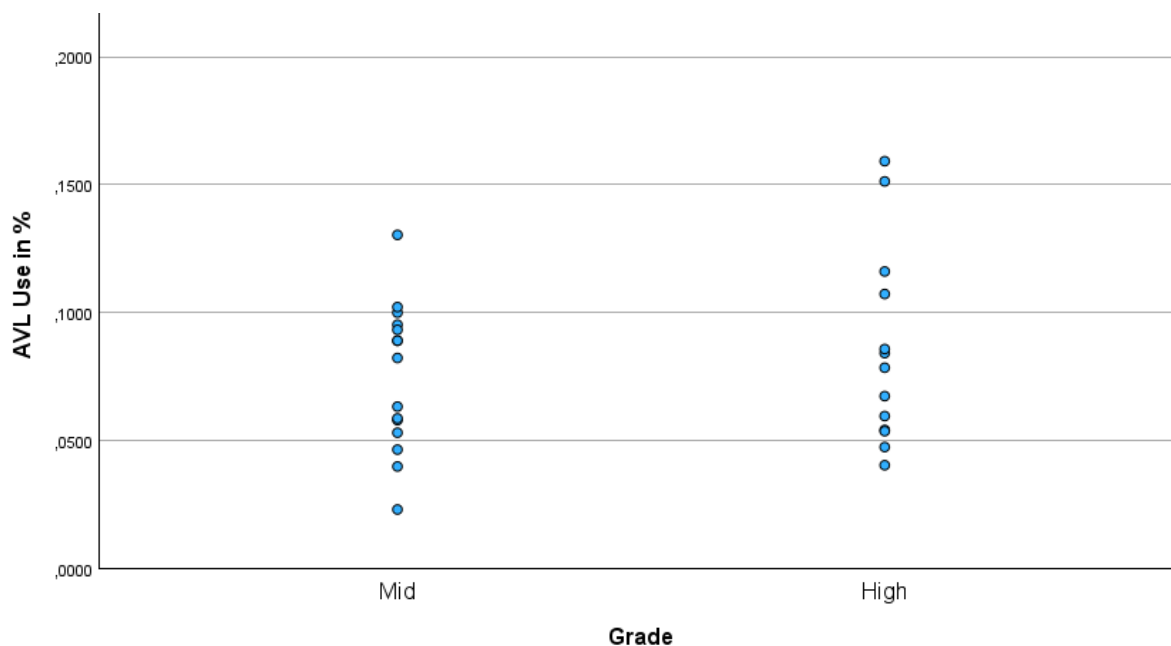
*AVL Use in Percentages in the Two Groups Based on Grades*

Grade	N	Mean	SD	Median	Min	Max	Range	Q1	Q3
Mid	15	7.49	2.87	8.23	2.31	13.04	10.73	5.55	9.43
High	13	8.50	3.84	7.85	4.04	15.92	11.88	5.42	10.73

*Note.* Q = quartile.

Table 13 shows that the high group produced general academic vocabulary items from the AVL to a slightly higher extent, on average, compared to the mid group. The range in values is 10.73 and 11.88 for the mid and high group, respectively. In addition, the standard deviation is higher for the high group (SD = 3.84) than for the mid group (SD = 2.87). Thus, there is slightly more variation within the high group.

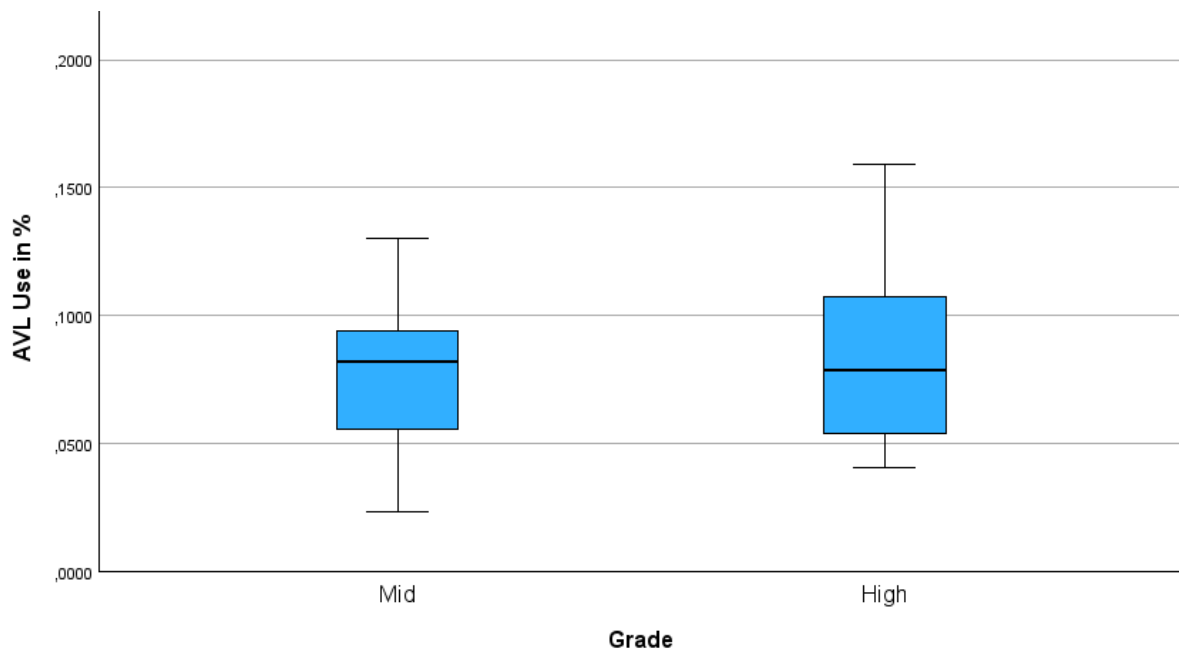
Graphic displays of this data were created in SPSS. First, a scatter plot was created to illustrate the spread in AVL production, shown in Figure 3

**Figure 3***Scatter Plot of AVL Use by Grade*

*Note.* In the dataset, percentages were written as decimals, thus, .1500 on the y-axis is 15%.

Larson-Hall and Plonsky (2015) describe a scatter plot as an “excellent data-accountable graphic” because it lets “the viewer ... see all of the variability and uncertainty in the data before their eyes, and thus has a rubric against which to judge the author’s interpretation of the data” (p. 147). From the scatter plot, we can see how the individual values in each group are dispersed. This dispersion, or variability in values, is somewhat similar for both groups. This indicates a low degree of consistency in AVL production on the written English mock exam among learners irrespective of achieved grade. However, with a 3.84 standard deviation score, the AVL vocabulary use among the learners in the high group is more dispersed than the distribution in the mid group. With a 2.87 standard deviation score, the learners in the mid group produce AVL vocabulary in their mock exams in a more homogenous way. Nonetheless, the group with the higher grades is slightly above the mid group on the y-axis, overall. Both the minimum and maximum AVL production values, as well as the mean, are larger in the high group compared to the mid group. Thus, we see tendencies towards a positive relationship between AVL use and achieved grades, but only a modest one.

In addition, a boxplot was created in the same statistical software tool. This is visualized in Figure 4.

**Figure 4***Boxplot of AVL Use by Grade*

Boxplots are useful in that they allow us to “compare some distribution across groups” (Hudson, 2015). These graphics allow such comparison “without making any assumptions regarding the underlying probability distribution, but at the same time indicating the degree of dispersion, skewness, and outliers in the given data set” (Larson-Hall & Herrington, 2009, p. 370). As we can see from the length of each boxplot and the corresponding whiskers, the total variation is somewhat similar. Moreover, the interquartile range (IQR) is slightly wider in the high group, which means that the values within the middle 50% in this group are somewhat more dispersed than in the mid group. Interestingly, the median is slightly larger in the mid group compared to the high group. However, there is a skewness in both distributions. On the one hand, the distribution in the high group is positively skewed, indicated by having a shorter whisker below the lower hinge (lower quartile) and a longer whisker above the upper hinge (upper quartile). On the other hand, the distribution in the mid group is negatively skewed, indicated by the median being closer to the upper hinge. This skewness likely explains why the median AVL vocabulary production in the mid group is greater than in the high group.

Further, potential differences in the reliance on general academic vocabulary in the assignment briefs and preparation material, between the two groups, were examined (see Appendix O). On the one hand, 65.22% of the AVL tokens produced by the mid group are



achieved by AVL types found in the assignment briefs, whereas this percentage increases to 78.98% when the preparation material is included. For the high group, the percentage of AVL tokens achieved by types found in the assignment briefs alone, or in combination with the preparation material, are 61.33% and 74.96%, respectively. Thus, while remarkably similar, the findings show that the mid group produces general academic vocabulary found in the texts in the assignment briefs and in the preparation material, to a slightly greater extent than the high group. Nonetheless, the relatively small differences possibly reflect the fact that both groups achieved considerably high grades.

Finally, using the first 500 words on the AVL word list to profile the AVL types and tokens produced by the learners in the respective groups, revealed small differences (see Appendix P). In the mid group, the top 500 lemmas on the AVL make up 83.82% of AVL tokens which, on the whole, cover 6.16% of the texts written by these learners. In the high group, the top 500 lemmas on the AVL make up 80.82% of AVL tokens, which overall cover 6.90% of the texts written by these learners. Thus, while these 500 most frequent words on the AVL make up a larger proportion of the general academic vocabulary tokens produced by the learners in the mid group, they provide a higher coverage of the texts written by the learners in the high group. However, the differences are only modest.

## 5.0 Discussion of Results

While the limitations of the present study are discussed later, it is worth to keep in mind the generalizability of the findings, in the following sections where the results are discussed. Findings should be discussed and interpreted within the boundaries of the purpose and scope which delimits the study (Durrant et al., 2022). It is therefore important to remember that this study examined 28 Norwegian L2 learners of English, in a general studies programme in upper secondary education, selected non-randomly, avoiding certain text-selection biases, but not all.

### 5.1 General Academic Vocabulary Production in the Corpus

#### 5.1.1 AVL Vocabulary Coverage

Gardner and Davies' (2014) AVL has not been used extensively in vocabulary research in general, or in vocabulary profiling in particular, due to its recency. Nonetheless, multiple studies have been conducted, which allows for comparison with the results in the present study (see section 2.6). However, it is important to be transparent about what separates these studies, in particular the size and nature of the corpora. Different ways of reporting results also make such comparisons less straightforward. Some studies report the percentage of general academic vocabulary in the entire corpus.

For instance, Malmström et al. (2016) reported an overall AVL token coverage of 19.3%, when profiling their Master of Science (MSc) corpus of approximately 720,000 tokens, written by first and second year master students. In Gardner and Davies' (2014) case study, the 570 AVL word families covered 13.7% and 13.8% of the academic sections of the BNC and COCA, respectively. In the newspaper sections of the same corpora, the coverage was 7-8%, and 3.4% in the fiction sections. Token size in these sections range from 10 to 120 million. For comparison, the AVL coverage of the 31,109 token corpus in the present study, is 7.91%. In other words, this is comparable to the amount of general academic vocabulary found in newspapers (Gardner & Davies, 2014). It is not unexpected that the proportion of AVL words is lower in texts written by learners in their first year of upper secondary school, compared to texts written by university master students (Malmström et al., 2016) and published academic writers (Gardner & Davies, 2014). However, these results are not entirely comparable, considering that Gardner and Davies (2014) used the 570 word family version of the AVL rather than the entire lemma based version, such as in the current study and the study by Malmström et al. (2016).

Other studies report not the entire AVL coverage, but the mean coverage of corpus texts. For instance, Durrant's (2016) profiling of the BAWE corpus of successful university student writing, revealed a mean AVL coverage of 16.82% (SD = 4.87) of all tokens. Again, this is significantly higher than in the present study, where the AVL, on average, provided a 7.96% coverage (SD = 3.33) of each written English mock exam. However, the results in the current study are closer to those found by Csomay (2020), where she profiled EFL student writing from a STEM field in a Georgian university. Her results showed that the students produced, on average, 9.94% AVL words (SD = 2.80). Further, in Csomay and Prades' (2018) study of ESL university student writing, the mean AVL vocabulary production ranged from 5.98% (SD = 1.68) to 12.06% (SD = 2.22). However, this was not reported for each participant. Instead, these measures differed based on variables such as instruction level, draft and text type. For comparison, when profiling texts written for each of the three tasks separately in the present study, the mean AVL coverage was 8.07% (SD = 3.24), 5.76% (SD = 2.52), and 8.76% (SD = 4.95) for task 1-3 respectively. Again, it is unsurprising that university students use general academic vocabulary in their writing to a greater extent, as Durrant's (2016) findings illustrate. Moreover, as mentioned in section 2.6.1, the corpus profiled by Durrant (2016) consisted of texts written by students in British universities, not specifically L2 learners of English. This could also explain the comparatively lower AVL use in the current study.

However, interestingly, the average AVL coverage of the texts in the present study is closer to Csomay and Prades' (2018) and Csomay's (2020) findings, both studies of university student writing. Nonetheless, even in these studies, the AVL coverage is greater than in the current study. Similar to how Malmström et al. (2018) found that university students' productive vocabulary is smaller and less academic compared to published writers, it is reasonable to expect that upper secondary learners' productive general academic vocabulary is even more limited.

Olsson's (2015) study of AVL vocabulary production among Swedish upper secondary learners, allows for a closer comparison with the present study. She found that AVL coverage ranged from 5.85% (SD = 2.38) to 12.14% (SD = 3.51), with variation based on instruction method and task assignment. However, only the essays written in response to the first assignment were written during the first year of upper secondary, similar to the learners in the present study. In the first assignment, both the mean AVL coverage for CLIL students at 7.37% (SD = 2.51) and for non-CLIL students at 5.85% (SD = 2.38), are lower than in the mock exams in the present study. Each mock exam in my corpus has a mean AVL

coverage of 7.96% (SD = 3.33). Thus, the average AVL use is slightly higher in the current study compared to Olsson's (2015), but so is the variation, shown by the larger standard deviation. However, even when considering only the texts in assignment 1 in Olsson's (2015) study, her corpus size is still significantly larger, containing 94 essays by CLIL students with a mean length of 570 words (SD = 207) and 52 essays by non-CLIL students with a mean length of 364 words (SD = 164).<sup>9</sup> For comparison, the corpus in the current study consists of 28 mock exams with a mean length of 1,111 words (SD = 143). This is important to keep in mind when comparing the results, as size is a key factor which impacts the representativeness of a corpus (see section 3.3.1.1).

### 5.1.2 Variation in AVL Use

It is interesting to note the variation in AVL vocabulary production among the participants in the present study. As the analysis of AVL coverage based on task revealed, AVL vocabulary use among the learners was most heterogeneous in the texts written for task 3 (SD = 4.95). The variation manifest in the present study might reflect the points Corson (1995, 1997) made on the lexical bar, and the impact of each individual's exposure to academic language in everyday life. The variety in AVL production could perhaps indicate that encountering academic vocabulary in English discourses outside of school is highly individual, leading to idiosyncratic reliance on such lexis in productive use. Indeed, Olsson and Sylvén (2015) suggest that the Swedish upper secondary learners in their study might rarely encounter general academic vocabulary when they use English outside of school.

From a usage-based perspective, the acquisition of general academic vocabulary is contingent on exposure to such vocabulary in the input, in order for it to be entrenched and ready for use (see section 2.1). In fact, within a usage-based approach, the concept of variation is central. Considering that "individuals will not have exactly the same experience in life, their development will not be exactly alike" (Verspoor & Smiskova, 2012, p. 22) Given the emphasis which is placed on experience, either through exposure or practice, Verspoor and Smiskova (2012) postulate that variation, from a usage-based perspective, can be regarded as "a normality rather than an aberration" (p. 22). In addition to differences in experience, individual variation in general academic vocabulary knowledge may also be attributed to other factors, such as aptitude (Webb & Nation, 2017).

---

<sup>9</sup> For an overview of the corpus composition in Olsson's (2015) study, see Appendix A on page 72.

However, the fact that learners were given the choice between a number of different prompts in task 3, suggests that the variation in AVL production might also reflect the different tasks and topics. The impact of topic on AVL use in student writing is highlighted by Therova (2022), who suggests that “some topics may prompt the higher usage of academic vocabulary” (p. 135). On the one hand, it could be argued that the production of general academic vocabulary should be minimally impacted by the particular tasks prompting the written texts. Indeed, the very definition of general academic vocabulary is that this lexis is not specific for any domain or subject (Baumann & Graves, 2010; Nagy & Townsend, 2012). On the other hand, task choice is likely to affect vocabulary production in general (Read, 2000), something which has been shown to be the case for AVL production, as well (Olsson, 2015). This might have had an impact in the present study as well. Nonetheless, with a standard deviation of 3.33 for each mock exam, AVL vocabulary use among the learners in the present study is not entirely homogenous, overall. Thus, variation in AVL vocabulary production is most significant in, but not limited to, the texts written for task 3.

### **5.1.3 AVL Vocabulary from the Top 500 Words on the List**

Moreover, profiling my corpus using the first 500 words on the AVL, revealed that these words make up 82.28% of the entire AVL coverage. These results are comparable to those reported by Malmström et al. (2016) who found that the top 500 words on the AVL made up 70.5% of AVL tokens produced in their corpus, and who compared this to the 74% coverage of AVL tokens in the COCA academic section. It is also somewhat comparable to Durrant’s (2016) study, who found that 313 words on the AVL accounted for two-thirds of the AVL vocabulary coverage for all the lexical words. The higher production of the first 500 words on the AVL in my corpus might be expected from learners in the first year of upper secondary school compared to university students and academics. Indeed, as the results of Malmström et al. (2018) indicate, university students use more high-frequency academic words compared to published writers (see section 2.5.2). Therefore, we might assume that the more infrequent words, further down the list, can be more difficult to use. From a usage-based perspective, frequency is an essential factor in the acquisition of vocabulary, and less frequent words might be less salient in the input, and less likely to be noticed and acquired (N. Ellis, 2013; Loewen & Reinders, 2011). As discussed in section 2.3.1, infrequent words might also contribute to making the gap between receptive and productive knowledge, wider (Webb, 2008). This suggests that besides acquiring more infrequent general academic words, they might also be more difficult to produce.

However, the fact that the word *value* is on this list, may be problematic, given that this is the most frequently produced AVL word in the corpus. As we can see from Appendix E, this word is among the words which appear on the AVL in more than one word class. While the noun lemma *value* is listed as number 33 on the word list, the verb lemma *value* is listed as number 809. Thus, only the noun lemma is part of the top 500 words. This is a limitation connected to the use of an untagged corpus, discussed in section 3.4.1.2. Of the 50 words from Appendix E which are produced in my corpus (see Appendix F), 19 of them appear in one POS within the top 500, and in another POS outside the top 500. Thus, the higher proportion of general academic vocabulary which is made up of the first 500 words on the AVL in the present study compared to other studies, might be attributed to the use of an untagged corpus. This limitation notwithstanding, the results from this analysis suggest that the learners mostly produce AVL tokens from the first 500 words on the AVL.

The amount of general academic vocabulary tokens covered by the top 500 words on the AVL might suggest that the word list is too extensive. Yet, the fact that a small number of AVL words are responsible for a significant portion of the coverage, is not entirely unexpected considering Zipf's law (see section 2.5). In addition to the 82.28% AVL coverage provided by the first 500 words on the AVL, 51.04% of the AVL coverage is achieved by the 50 most frequently produced AVL word types in the corpus (see Appendix G). These results support the findings in Durrant's (2016) study, namely that a small portion of words account for a substantial number of tokens, although in the present study, this skewness is even more apparent. On the whole, this points towards similar tendencies, indicated by Durrant (2016) and Malmström et al. (2018), namely that productive academic vocabulary is somewhat limited.

#### **5.1.4 The non-AVL vocabulary**

Considering that 7.91% of the corpus tokens were AVL items, we might instinctively assume that the remaining 92.09% of the tokens are general, high-frequency vocabulary. However, some of these tokens are proper nouns. Profiling the corpus with the BNC/COCA lists (Nation, 2017) as reference lists, reveals that 98 types and 488 tokens are covered by basewrd31, the list containing proper nouns. This makes up 1.57% of the corpus tokens. In addition, the most frequent words in any corpus of written English are grammatical or function words (Oakey, 2022). These words likely make up a considerable portion, which is important to keep in mind. In fact, after lemmatizing the corpus, *the* ( $n = 1640$ ) and *be* ( $n = 1463$ ) alone make up 9.97% of the corpus. Together with the following eight most frequent

words *to* ( $n = 936$ ), *and* ( $n = 931$ ), *a* ( $n = 642$ ), *of* ( $n = 630$ ), *that* ( $n = 546$ ), *in* ( $n = 509$ ), *have* ( $n = 412$ ), and *we* ( $n = 387$ ), these ten words combined make up 26.02% of the corpus. Thus, if such words were excluded from the profiling, it would have revealed a much higher AVL percentage. Indeed, as Durrant (2016) demonstrated, excluding grammatical words led to a doubling in the mean AVL coverage, from 16.82% (SD = 4.87) to 33.82% (SD = 8.40). Nonetheless, it is most common in the literature and research to refer to coverage of all running words (tokens). Hence, this was the approach adopted in the present study.

Moreover, as pointed out by Csomay (2020), the non-AVL words produced by students may not necessarily be general words. Instead, it might be technical, subject specific vocabulary, which is more specialized, perhaps contributing to “a more sophisticated discussion of a given topic” (Csomay, 2020, p. 24). Further investigation would be required, however, in order to make any claims about the presence of such vocabulary in the corpus composed for the current study. However, given the large amount of coverage which has been shown to be provided by high-frequency, general words (e.g. Gilner, 2011), we may safely ascertain that such lexis indeed makes up a considerable portion of the non-AVL vocabulary.

Yet, while often considered in contrast to academic vocabulary, Granger and Larsson (2021) illustrate that this might not essentially be the case. The authors found that while general, or basic, vocabulary may indeed appear informal from a single-word approach, such lexis might be part of multi-word constructions regularly produced in academic texts. Nonetheless, the authors only investigated the lemma *thing*. However, Durrant (2009) found similar indications, namely that general vocabulary are often part of academic collocations. The possibility that some of the non-AVL words produced in the corpus might appear academic from a multi-word approach, is therefore worth bearing in mind. Yet, it might be equally probable that even from such a perspective, the general vocabulary only appear general.

### **5.1.5 Potential for Development**

The first year of upper secondary education in Norway is the last obligatory year with English instruction for learners in general studies. These learners should therefore be able to understand and produce general academic words in English to an extent which will allow for proper participation in tertiary education discourse. As the results in the present study revealed, the learners produced general academic words to an extent similar to what might be found in newspapers (Gardner & Davies, 2014). In other words, much learning potential pertaining to the productive use of this lexis, still exists. It is worth to repeat the words of

Gardner and Davies (2016), namely that “the absence of sophisticated academic vocabulary in the compositions of developing writers is not evidence that such words are unimportant now or at some future date for those writers” (p. 63). Moreover, it is by no means expected that the productive general academic vocabulary knowledge among the learners in the present study is final at the point of completing the first year of upper secondary school.

Based on results from previous research, it is reasonable to expect that L2 learners’ productive knowledge of general academic vocabulary might increase with time. In Olsson’s (2015) study, the average AVL production increased for all participants (CLIL and non-CLIL), from the first to the fourth assignment. These were written during the first and last year of upper secondary school, respectively, which suggest that some progress might be possible even before reaching university. Further, Csomay’s (2020) results indicate that AVL use increases as university students advance in their studies. Similarly, AVL use was marginally higher for university students in their second master’s year compared to the first, in the study by Malmström et al. (2016).

Conversely, Csomay and Prades’ (2018) results showed that students in the lower-level class produced AVL words to a greater extent than the students in the upper-level class. This indicates that a higher instructional level does not automatically lead to an increased production of general academic vocabulary. However, it should be noted that in Csomay and Prades’ (2018) study, the lower-level class was taken by everyone, whereas the upper-level class was only taken by students who failed a writing proficiency exam. Thus, as the authors point out regarding the students in the upper-level class, “their proficiency level may in fact be lower” (Csomay & Prades, 2018, p. 114). Yet, other studies also indicate low levels of sophistication in the productive vocabulary use among learners. Indeed, Laufer (1998) found that even as high school students increased their vocabulary knowledge, they did not use these words in free written production. As discussed in section 2.3.2, learners might avoid the risk of trying out more advanced lexis, and use more general words that manages to communicate meaning in their written production (Laufer, 1998). Indeed, the safety in using general rather than more advanced lexis, is that they can be widely used without being underlined as errors (Hasselgren, 1994). Thus, the question of whether learners actually progress in their productive use of advanced words, such as general academic vocabulary, is not a given.

Finally, irrespective of whether learners progress in their general academic vocabulary knowledge from upper secondary to university, their productive vocabulary might still differ significantly from the words on the AVL (Malmström et al., 2018). Thus, a question that



remains is whether the AVL should be considered the norm for learners aiming for university, or if other lists such as the PAVL (Malmström et al., 2018) better fulfill this purpose.

## 5.2 Impact of Assignment Briefs and Preparation Material

The investigation of a potential link between the AVL words produced by the learners, and the vocabulary in the assignment briefs and preparation material, revealed extensive overlap. In the present study, more than half (63.23%) of the AVL tokens produced in the corpus of learner texts are made up of types found in the assignment briefs. Further, three quarters (76.92%) of all the AVL tokens are made up of AVL types found in the assignment briefs and preparation material, combined.

These results corroborate the findings of Therova (2021, 2022) and Milton (2001). As Milton (2001) found, students tended to mirror the words in the prompts. In addition, Therova (2022), found that a large portion of AVL words used in student writing were from the assignment brief. In her study, however, the proportion was significantly lower than in the present study. She found that 14.36% and 22.86% of AVL tokens in two essays were achieved by types from the assignment briefs. This difference between the current study and Therova (2022) might possibly be attributed to the considerably larger corpus size in her study. Her corpus consisted of 386,439 tokens, while my corpus was 31,109 tokens in size. Another possible explanation might be that the tasks in the present study were related to topics that coincidentally happened to be AVL words. For instance, *change* and *value* were among the most frequently produced AVL words in the present corpus and were part of the overall mock exam topic in one school. Along similar lines, Therova's (2021) study revealed that many of the most frequently produced AVL words were connected to the assignment topic.

Furthermore, the results in the present study support the findings of Leki and Carson (1997), who suggest that students use source texts to find appropriate vocabulary. The findings by Plakans and Gebril (2012), who found that students use source texts for language support, such as vocabulary, are also corroborated. Results from the current study suggest that learners in upper secondary school use texts in the writing process in similar ways. On the whole, this points to the fact that there is a relationship between reading and writing, underlined by Hirvela (2016) and Coxhead and Byrd (2007). Moreover, given that one of the source texts which I considered part of the assignment briefs was a video, this relationship

might include the receptive skill of listening, as well as reading. Indeed, Schmitt (2014) claims that “virtually all aspects of vocabulary knowledge seem interrelated” (p. 942).

However, considering that the vast majority of productive general academic vocabulary was found in the assignment briefs and preparation material, it might be difficult to establish how much of this vocabulary has been acquired by the learners, versus how much is mirrored from the texts at their disposal during the writing process. Indeed, as outlined in section 2.3.2, writing is a difficult skill, requiring knowledge of spelling, collocations, and register, among others (N. Ellis & Beaton, 1993; Laufer, 2005; Nation, 2022). Writing formal texts might be considered even more challenging (Laufer, 2013). Relying on vocabulary available to the learners in the assignment briefs or preparation material, might be a way to compensate for this difficulty associated with vocabulary knowledge for written use. It might also lessen the burden of finding appropriate lexis to implement in the learners’ writing (Leki & Carson, 1997).

Nonetheless, the implementation of topic related vocabulary is also a way to follow the instructions on the mock exam tasks properly. It is not only a mirroring or borrowing of vocabulary in the assignment briefs or preparation material, but a necessary implication of discussing a given topic. Similar remarks are made by Therova (2022), who suggests that integrating general academic vocabulary from the assignment briefs is a way to address the instructions of an assignment. Moreover, from a usage-based perspective, vocabulary is learned through experience, both by exposure to input and practice in output (see section 2.1). Thus, by reading AVL words in the preparation material or assignment briefs, and subsequently producing these words when writing the mock exams, learners in the present study might have increased their knowledge of this lexis.

To some extent, these findings contradict the way in which general academic vocabulary is often defined, as more detached and less central to topic (Coxhead, 2000; Nation, 2022). However, it might also be true that the most salient and noticeable general academic vocabulary is that which is related to the topic, such as *change* and *value*. Thus, it might be that many of the general academic words less detached from the topic and content of the texts, have not been acquired due to their low salience. On the whole, however, these results support the findings by Therova (2021) who found that her case study participant used many topic related AVL words. She argues that this blurs the distinction between technical and academic vocabulary and suggests that there might be an intersection between the two (Therova, 2021). The former is defined as common to certain disciplines and their content,

whereas the latter is defined as common between disciplines, without being “strongly attached to any one particular subject area” (Nation, 2022, p. 275).

Yet, the influence of topic on vocabulary use is not completely unexpected, as this has been emphasized (Nation, 2022; Therova, 2021, 2022). In addition, the link between topic and vocabulary might also reflect the failure to avoid the topic bias when creating the corpus of learner texts (Brezina, 2018). In a way, it also confirms that the topic bias is real.

### **5.3 Correct Use of General Academic Vocabulary**

First of all, the most striking result from the qualitative analysis evaluating each AVL token produced in the corpus, is how few instances were categorized as incorrect. Merely 20 tokens were evaluated as being incorrect due to an effect of semantic dissonance, out of 2,461 AVL tokens in total. Moreover, about half of the participants' mock exams (46.43%) did not consist of any incorrect instances, at all. This indicates that once general academic vocabulary is acquired, the learners use them correctly in terms of their meaning. Still, it is worth underlining yet again that a lenient approach to determining correct use was taken in this study. A complete error analysis of all AVL words in the corpus would perhaps reveal instances of words being used incorrectly in terms of stylistic or collocational effects, or along other lines (Hasselgren, 1994). However, such an analysis was beyond the scope of this study and should ideally be based on a larger number of AVL tokens.

In addition, the corpus was not corrected for spelling mistakes. Such an analysis could have had two potential outcomes. First, it might have added some words to the number of AVL items present in the corpus, which could increase the overall AVL coverage. Second, it could provide insight into the extent of AVL words with orthographic errors, as well as which types this applies to. However, this was not within the scope of the current study, as it would have required manually spell-checking all tokens in the corpus, before the AVL profiling. While possible, it would be time-consuming. However, as Webb (2008) established, productive vocabulary knowledge among the learners in his study was higher when words spelled incorrectly were included. This could be the case for productive general academic words, as well.

Nonetheless, it is interesting to observe that several of the AVL words deemed incorrect appear as having been misspelled. Consider the verb lemma *effect* in the following examples:

- (1) The decision whether it is good to change the past is complex as it **effects** people from all sides.
- (2) The values of a person will have great impact on the persons decisions, and therefore **effect** one's life.

In both cases, it would be reasonable to assume that a simple spelling mistake was the cause of the error, given that the verb lemma *affect* is the appropriate word in both contexts. It might suggest that the learners have heard the word more frequently than seen it, considering the phonological similarity when pronounced. However, the similarity is present in the written mode as well.

Laufer's (1988) taxonomy of synforms may help explain some of the detected errors (see section 2.1.2). One of these categories, number six, includes words which are identical except for one vowel, and *affect/effect* is in fact used to exemplify this (Laufer, 1988). Other categories also help elucidate incorrect use of general academic vocabular in my corpus. The third category in her taxonomy includes words similar in form, but one of the words contain a suffix that the other does not have. For instance, in the following example, a suffix is the only element separating the produced word, *previous*, from the intended one, *previously*:

- (3) Lastly, one should also make sure one does not recycle **previous** submitted works, as this is seen as self-plagiarism.

Further of Laufer's (1988) categories, such as the first one, describes words identical in root but differing in suffix. One general academic word from my corpus which may be included in this category is highlighted in the following example:

- (4) Values are often passed down through family, **development** from own experiences as well as history.

In this example, *development* differs from the intended meaning *developed* only with regard to the suffix in each respective word. Moreover, consider the mistake of confusing *archives* with *achieves* in the following sentence:

- (5) This is proven by looking at the old **achieves**.

This error might not fit in any category in Laufer's (1988) taxonomy, but a similarity between their written forms is nonetheless evident.

However, these mistakes could also be due to a lack of experience with general academic vocabulary, on the whole. Corson (1997) posits that people whose experience with academic language is limited may not be able to process or activate the appropriate academic Graeco-Latin word. One consequence of this, he argues, is that "a morphologically similar but inappropriate word might be activated and used in the message (e.g., *concept* for *context*), an irregularity that we usually put down to "a slip of the tongue.'" (Corson, 1997, p. 699, emphasis in original). This can explain the mistakes exemplified above. However, other errors might be more in line with such an explanation. For instance when *provide* is produced instead of the appropriate word *prevent*, in the following sentence:

- (6) That is why it is super important to be well educated about these kinds of things, so that we are able to **provide** and avoid them from happening again.

All of these examples illustrate a potential difficulty that may arise when making form-meaning links in the L2, which is how language is learned from a usage-based perspective (Verspoor & Schmitt, 2013). When the forms of lexical items are similar, it might be difficult to distinguish them, which can make them challenging to learn for L2 learners (Verspoor & Schmitt, 2013). Furthermore, if it is challenging to distinguish words in the input, they may not be considered particularly salient. At least not the features which differ, such as a vowel or suffix. Noticing the different ways that such words are used can therefore be difficult, and the learners might struggle with some of the aspects of productive vocabulary knowledge required for writing (see Table 2). The results from this analysis suggest that explicitly focusing on form might be beneficial in certain cases where the form of lexical items differ only marginally. On the whole, however, it seems as if the learners use general academic vocabulary correctly in terms of correct meaning.

When determining correct use, an additional insight was gained in the process. Words can be used semantically correct within the context of the sentence in which it is produced, and yet, the meaning may not be academic. Consider the following example, extracted from the corpus:

- (7) [T]hey are going to have their feet on the **table**.

In this sentence, the AVL word *table* is not incorrect in any sense. Moreover, it is used as a noun, the same POS as on the AVL. Still, it does not appear academic. It seems sound to assume that the word is academic in the sense of columns and rows presenting data or information, not the furniture. Coxhead (2020) attributes this to Gardner and Davies (2014) defining general academic vocabulary as words across all frequency levels. As a result, whether a word appears academic or general depends on the context or discipline (Coxhead, 2020). It would, however, be difficult to determine the extent to which each AVL word produced in the corpus is academic enough. Evaluating which meaning is the most academic for each AVL item, in cases where meaning variation is present, would have included multiple assumptions and subjective judgements.

Still, it is worth considering the existence of meaning variation, something Gardner (2007) highlights as a potential validity issue in corpus-based vocabulary research. One way to cope with the issue of semantic disparity would be to use a word list where the lexeme is used as the unit for counting words. However, as noted by Brezina (2018), this would involve “semantic tagging and semantic disambiguation, which again introduces a certain percentage of errors (even higher than part-of-speech, or POS, tagging) and, moreover, cannot be done fully automatically” (p. 41). While the extent of meaning variation in the present study was not examined, it has been used as an argument against the notion of a core general academic vocabulary (Hyland & Tse, 2007). Nation (2022) addresses this issue by arguing that it is crucial to recognize how both the core and specific meaning of academic vocabulary can be useful, instead of viewing the two as mutually exclusive.

Finally, in the process of manually searching for and removing instances of overcounting, one could argue that this contributed to the process of evaluating correctness, as well. Certain words which would have been deemed incorrect, were excluded from the AVL count because their POS differed from the POS for the word listed on the AVL. One example includes using the word *overview* as a verb where *overlook* was the intended meaning. *Overview* is an AVL word exclusively as a noun and was thus disregarded as an academic word. Another example is the word *precise*, appearing on the AVL as an adjective, but being used in the corpus as a verb with the intended meaning of *specify*. Additionally, the word *continuous* is an adjective but was used as a verb on more than one occasion, in the sense of *continues*. This was also excluded from the AVL count because of being the wrong POS. Therefore, the removal of these words and other similar ones aided in strengthening validity, primarily intended to be accomplished through evaluating correct usage.

#### 5.4 Relationship with Achieved Grades

From the graphical visualizations of AVL production among two groups of learners based on grades, we can see a tendency that learners in the high group produced AVL words to a slightly greater extent than the learners in the mid group (see Figure 3 and Figure 4). In other words, these indications point to a positive relationship between the use of general academic vocabulary and achieved grade. However, the differences in AVL production between the two groups are not particularly large. Additionally, while the high group had a slightly higher average AVL production, the mid group had a slightly higher median AVL production. Perhaps of greater interest is the fact that the use of general academic vocabulary was considerably dispersed within both groups.

Csomy and Prades (2018) did not find a statistically significant correlation between essay score and the productive use of AVL vocabulary, overall, only in some text-types and drafts. The authors attribute the lack of a systematic relationship between the production of general academic vocabulary and achieved grades to variation in text types, among other. This could potentially be a factor impacting the grades in the current study as well, given that texts elicited by three different tasks were included. Furthermore, in Olsson's (2015) study, essays evaluated as weaker and stronger than the essay used for comparison had a mean AVL vocabulary coverage of 10.57% (SD = 3.37) and 13.66% (SD = 3.67), respectively (see section 2.6.3). In the present study, participants in the mid and high group produced an average of 7.49% (SD = 2.87) and 8.50% (SD = 3.84), respectively. This indicates similar tendencies, namely that AVL production positively impacts assessment, corroborating the findings of Olsson (2015).

Further, Nagy and Townsend (2012) maintain that improving learners' achievement is the primary reason for supporting their proficiency in academic language. Considering the link between academic language proficiency and academic success, they suggest expanding assessment of student writing beyond academic vocabulary only. The authors highlight some of the other characteristics of academic language, such as morphologically complex words, prepositional phrases, and nominalizations, among others. It is not unlikely that other elements making up academic language, such as these, contribute to the holistic assessment of student writing. Other aspects are also likely to influence the impression of how academic a written text is, such as the topic or audience, as well as the communicative purpose (Dixon et al., 2023). Moreover, in the process of assessing student writing in any subject, disciplinary content knowledge is also likely to be of importance. It is for this reason that Nagy and Townsend (2012) argue that both "disciplinary knowledge measures and academic language

measures, are proxies for a student's overall academic language proficiency" (p. 104). Thus, the teachers who graded the written English mock exam texts making up the corpus in the current study, may have considered a variety of criteria and factors in their grading process.

In addition, it is necessary to mention that the grade each participant achieved on their mock exam is likely to have been impacted by the results from the digital, multiple-choice task. This was the first out of four tasks making up the mock exams and was not included in the present study. Given that the four parts in the mock exam were not assessed individually, but as a whole, not knowing the results of the multiple-choice task is one limitation.

Finally, the reliance on a small set of AVL words, the first 500 on the word list, was similar for both groups. Moreover, the reliance on AVL words from the assignment briefs and preparation material was strikingly similar, as well. This suggests that learners irrespective of grades rely on a small set of words from the AVL, and that a considerable portion of this is from texts they read in the writing process, including the prompts themselves. However, all the grades were relatively high, with no grades being lower than 4-. A more diverse sample of participants might have revealed greater variation.



## 6.0 Conclusion

The next section highlights the key findings from the current study, which was guided by the following three RQs:

1. To what extent do Norwegian L2 learners of English in upper secondary general studies produce general academic vocabulary in written mock exam texts?
  - a. To what extent is this vocabulary production influenced by the words found in the assignment briefs and preparation material for the mock exam?
2. To what extent do these learners use general academic vocabulary correctly?
3. Is there an association between the extent of general academic vocabulary production in the written texts and achieved grades for these learners?

General academic vocabulary was operationalized as words on the Academic Vocabulary List (AVL) by Gardner and Davies (2014). This word list was used to examine a corpus of 31,109 tokens comprised of English mock exam texts written by 28 Norwegian L2 learners of English in their first year of upper secondary education, in a general studies programme. The analyses of this data were mainly conducted in *AntWordProfiler* (Anthony, 2023).

### 6.1 Key findings

AVL vocabulary covered 7.91% of the tokens in the entire corpus. The average AVL vocabulary production for each learner was 7.96% (SD = 3.33). This is lower than the 14% which is typically found in academic texts across disciplines, but comparable to the 7-8% found in newspapers (Gardner & Davies, 2014). However, AVL vocabulary use was heterogeneous, which means that some learners used this lexis to a greater extent than others. Learners who achieved the highest grades produced, on average, slightly more AVL vocabulary than learners with a lower grade. Nonetheless, only descriptive statistics were used, and distributions were considerably dispersed irrespective of achieved grades. Thus, there was wide variation overall in the extent to which the learners used general academic words in their written texts.

While the AVL consists of 3,015 words (lemmas), the first 500 words on the list covered the majority (82.28%) of the AVL tokens produced by the learners in my corpus. This is comparable to other studies, where this part of the list has shown to cover 70.5-74% of AVL tokens in various corpora (Malmström et al., 2016). In addition, the 50 AVL types with

the highest frequency in my corpus made up 51.04% of the AVL vocabulary tokens produced by the learners in the present study.

Furthermore, most of the AVL tokens produced in the corpus, 63.23%, were made up of AVL types present in the assignment briefs. If we include the preparation material, this number increases to 76.92%. Thus, most of the general academic vocabulary coverage in the corpus is provided by only a handful of word types, primarily from a small portion of the AVL, and overall, most of this vocabulary is found in the assignment briefs and preparation material.

Finally, analyzing the AVL tokens in the corpus for correct use revealed that only a marginal number of words were used incorrectly. However, the only criteria guiding this evaluation was semantic dissonance. Nonetheless, this indicates that the learners in the present study mostly produced general academic vocabulary correctly in the sense that these words did not lead to wrongness in meaning. Yet, the instances that were deemed incorrect, suggest that similarity in form between the intended and produced word might be a factor impeding the correct use of general academic words.

The findings in this study contribute with insight into the written use of general academic vocabulary by Norwegian L2 learners of English in upper secondary. I hope the findings can contribute to a better understanding of how learners use this lexis in high-stakes writing such as the mock exam. Overall, this study provides a foundation that further studies with a productive focus on general academic vocabulary can build on or use for comparison. The present study also demonstrates the variety of approaches it is possible to take when analyzing written texts for general academic vocabulary. This includes to zoom in on particular words and evaluate their correctness, or to zoom out by examining the potential impact of reading material and assignment briefs on the vocabulary production. These decisions are, however, clearly contingent on corpus size and access to contextual information.

## **6.2 Implications**

### **6.2.1 Increased Focus on General Academic Vocabulary**

The wide variation in general academic vocabulary production among the L2 learners of English in the present study, indicate that an increased focus on this lexis can be beneficial. Many learners might not encounter general academic words in English outside of school (Corson, 1995, 1997; Olsson & Sylvén, 2015). Because experience with the language is how

it is acquired, from a usage-based perspective, teachers should provide learners with opportunities to encounter and use general academic vocabulary in the L2 classroom. Ideally, this should happen several times, and in a variety of contexts (Verspoor & Schmitt, 2013). This can include reading more academic texts in English, although this might not be sufficient by itself (Cons, 2012; Hinkel, 2003). Indeed, encountering these words can contribute to the acquisition process. However, the low perceptual salience of general academic words in the input can be an impeding factor (e.g. Corson, 1997; Coxhead, 2000; Nation, 2022). Thus, teachers can utilize different techniques which involve modifying the input, as a means to construct salience, making general academic vocabulary more noticeable (R. Ellis & Shintani, 2014). In addition, an increased focus on this lexis should involve more writing practice, where learners are given the opportunity to produce general academic words (Cons, 2012). Yet, explicit instruction of these words, and how they are used, might still be necessary.

Several researchers have called for the explicit instruction of general academic vocabulary (e.g. Csomay, 2020; Csomay & Prades, 2018). This has shown to have an effect on vocabulary use in L2 writing (Lee, 2003). As outlined in section 2.3.2, productive vocabulary knowledge is not necessarily a by-product of receptive vocabulary knowledge (Laufer, 1998; Laufer & Paribakht, 1998; Lee, 2003). Thus, for receptive words to become productive, explicit instruction might be necessary (Schmitt, 2008). Moreover, the fact that the learners mostly used general academic vocabulary from the first 500 words on the AVL, indicates that it might be beneficial to focus the learners' attention on the less frequent general academic words, as well. In addition, results from analyzing general academic vocabulary for correct use indicate that explicitly focusing on words similar in form might also be useful, as these can potentially cause confusion (Laufer, 1988). Overall, explicit focus on general academic words can be one way to increase the amount of experience L2 learners have with this lexis. This is one way for teachers to help lower the lexical bar (Corson, 1995, 1997), which can help make the process of transitioning into tertiary education more seamless, where both receptive and productive knowledge of English general academic vocabulary, is important (Paquot, 2010).

If teachers decide to explicitly focus on general academic vocabulary in the L2 classroom, it can be useful to monitor the learners' use and development of this lexis. This would give the teacher a more precise and complete picture of the learners' current state of productive general academic vocabulary knowledge. Such insight could be gained by studying their written texts, for instance (Hestetræet, 2020). Similar points are made by Webb and Nation (2017), who suggest administering tests measuring vocabulary knowledge. In

addition to tests, teachers might also rely on corpus software tools, such as the one utilized in the present study. Based on such insights, the teacher can evaluate the need to focus further on this vocabulary. Furthermore, when determining what lexis to monitor or focus on explicitly, teachers might use a word list of general academic vocabulary. The AVL can fulfil this purpose. However, it has been suggested that this list might be too extensive (e.g. Durrant, 2016; Malmström et al. 2018), and that it might be more suitable for research purposes, than in the L2 classroom (Webb & Nation, 2017).

Nonetheless general academic vocabulary is a kind of lexis where teachers can provide more help to learners of English, in contrast to more specialized, technical lexis (Nation, 2022). Therefore, Nation (2022) posits that an academic vocabulary list is worth focusing on, both for teachers and learners. However, teachers wanting to use such lists should have knowledge of the methodological criteria and objectives behind the word list (Coxhead, 2016a). Moreover, the notion of word lists in general has been problematized from a communicative perspective. It can be argued that it contradicts the assumption within a communicative approach, namely that L2 vocabulary develops from communicative exposure (Zimmerman, 1997). Nonetheless, Flowerdew (2012) underscores that corpora have influenced several aspects of pedagogy, in spite of doubts directed toward the relevance of corpus derived frequency lists, and corpora in general. The relationship between corpus linguistics and language teaching is, however, not entirely straightforward. In fact, Seidlhofer (2003) highlights this relationship as one of the controversies in the field of applied linguistics.

### **6.2.2 Vocabulary Learning Potential in Reading Texts when Writing**

Another implication of the current study is related to the link between general academic words produced by the learners, and the vocabulary in the various reading material the learners had at their disposal. This relationship between reading and writing has been emphasized by researchers (Coxhead & Byrd, 2007; Hirvela, 2016). Indeed, most of the AVL tokens produced by the learners were achieved by types found in the assignment briefs and preparation material. Hence, the findings in the current study suggests that learners take advantage of the vocabulary available in texts of various kinds, when writing. In the present study, this included a short video that they had to summarize. This material can therefore potentially serve as useful tools in eliciting specific general academic words.

Therefore, teachers can purposefully formulate the instructions in tasks and intentionally pick source texts to rely on or incorporate in the writing process, so that these texts include vocabulary they want to elicit from the learners. Similarly, when choosing what source texts to include in the writing process, if these are to be included, teachers should consider what vocabulary the texts might elicit. In this endeavor, teachers also have the opportunity to modify the input and make the target vocabulary more salient, for instance by boldfacing, underlining or italicizing words (R. Ellis & Shintani, 2014). Yet, even without such enhancement of particular words, the impact of reading texts on vocabulary production has been emphasized (Leki & Carson, 1997; Milton, 2001; Plakans & Gebril, 2012). This link has also been supported empirically for the written production of general academic words (Therova, 2021, 2022), and is corroborated in the current study.

Finally, L2 learners might use vocabulary from the texts available to them when writing, for several reasons. It might be a way to compensate for the lack of vocabulary knowledge or a consequence of following the instructions of a given task (see section 5.2). Nonetheless, by encountering general academic words in the input and producing them in the output, learners' knowledge of these words is likely to be strengthened, at least from a usage-based perspective (e.g. Verspoor & Smiskova, 2012).

## **6.3 Limitations**

### **6.3.1 Methodological Limitations**

Several limitations regarding the methods and analytical procedures in the current study were discussed in chapter 3. These included limitations linked to the use of an untagged corpus. Errors in the process of lemmatizing the corpus also introduced some difficulties (see section 3.4.1.2). Schmitt (2010) maintains that one must be careful with generalizing to an entire group of L2 learners. He argues that generalizations should be made along the lines of proficiency and L1, primarily. Given the central role of the L1 in interpreting results (Schmitt, 2010), it is a limitation that I did not gather information about this from the learners, and instead assumed that the L1 was Norwegian for all the participants. Other limitations are related to the failure to avoid certain biases in the text selection process, using a non-probability sampling technique, and having a small sample size (see section 3.2 and 3.3.1.3). These limitations impact the ability to generalize the results to the wider population.

Moreover, Szudarski (2018) identifies four limitations with corpus analysis when researching vocabulary. First, corpus data only consist of positive data, not all possible

features of language. Second, the corpus may not be large enough to sufficiently capture the language use we aim to describe. Third, corpus data is decontextualized. Fourth, no corpus can interpret the data for us. Thus, the findings and how accurate they are depend on the researcher, particularly for specialized corpora, since it is common that the same person both creates and analyzes such corpora (Szudarski, 2018). On the first point, it is true that the corpus only captures language that is part of the corpus, and nothing else. As Durrant et al. (2022) underline, it is likely that learners know more vocabulary than they use. Nevertheless, it does show what kind of words the learners choose to use freely, which is something separating this kind of research from studies using vocabulary tests, for instance.

Further, while it is true that no corpus can capture language in its entirety, the size depends on the aims of the research study (Reppen, 2022). Thus, by purposefully selecting data to describe a particular language feature, such as general academic vocabulary, a small corpus is unproblematic. However, principles such as representativeness, balance and size should guide the corpus creation, but is often determined by practical aspects, such as limited time, resources, or access to corpus material (see section 3.3.1.1) Moreover, the issue of context-stripping is not as prevalent when using small specialized corpora, such as in the current study (see section 3.3.1.2). Finally, the fact that corpora cannot interpret the data is a criticism that can be directed toward most types of research, not corpus linguistics in particular. Nonetheless, this limitation necessitates rigorous transparency about the decisions and steps in the research process.

### **6.3.2 Limitations with Single Words**

Finally, there is a limitation linked to the very construct under investigation in the present study, namely single-word units of general academic vocabulary. Focusing on isolated units is common in vocabulary research, a trend that Anthony (2020) describes as bizarre. However, Paquot and Plonsky (2017) consider this to be a serious methodological shortcoming in corpus linguistics, in general. A central criticism against academic vocabulary has been whether it is possible to have a word list containing single-item vocabulary, stripped of context. Indeed, Hyland and Tse (2007) argue that academic vocabulary is misleading as it does not take into consideration how its use varies based on discipline. In addition, the importance of considering context, collocations, and lexical bundles have been emphasized both when explaining how to use general academic vocabulary (Nagy & Townsend, 2012), as well as when describing academic discourse in general (Eldridge, 2008). A growing trend is to research patterns of multiple words as they occur together in discourse, such as lexical

bundles and collocations (Coxhead, 2016a). Some postulate that it would be beneficial to define vocabulary as formulaic vocabulary (Martinez & Schmitt, 2012).

However, the limitations associated with single-word units would not necessarily be mitigated by focusing on multi-word units, because the criticism is directed toward this as well. The issue is that these elements, whether single or multi-word units, are considered “the main diagnostic of importance” (Paquot & Plonksy, 2017, p. 63). It is therefore important not to overestimate the significance of single-word units of general academic vocabulary. Nonetheless, despite this limitation, single-word units of general academic words have been shown to make up a considerable proportion of tokens in English academic texts from various disciplines, as well as being linked to the assessment of written proficiency and academic success (Corson, 1997; Coxhead, 2000; Gardner & Davies, 2014; Nagy & Townsend, 2012). Thus, I would argue that it can still be of importance. However, this limitation underlines the importance of teaching single-word units in context, a point that has been emphasized extensively (e.g. Coxhead, 2000; Hyland & Tse, 2007; Nagy & Townsend, 2012).

#### **6.4 Recommendations for Further Research**

Considering the limitation of examining single-word units, it would be of interest to investigate the production of general academic multi-word units by Norwegian L2 learners of English. It is possible to profile written texts for the production of multi-word units, using much of the same methodology as in the present study. It is, however, more challenging to implement multi-word units in software tools providing coverage, which is a possible reason for why it has not received as much attention as single-word units in vocabulary research (Gablasova & Brezina, 2021). Nonetheless, various word lists consisting of academic multi-word units exist readily available for similar analyses. The need to focus on this in future research is acknowledged by the authors behind the creation of the AVL (Gardner & Davies, 2014). The authors also highlight the need to focus on spoken academic word lists. While there is an overwhelming tendency in studies of learner language to investigate written corpora, the mode used and analyzed in the current study, there has been an increase in the number of studies utilizing an oral corpus (Paquot & Plonsky, 2017). Thus, the investigation of general academic vocabulary in learner language from a multi-word approach, as well as in an oral corpus, would be interesting to focus on in future research.

Further investigation into the relationship between reading and the written production of general academic vocabulary, either from a single- or multi-word approach, would also be

interesting. Having access to contextual information, such as preparation material and assignment briefs, is an advantage with small, specialized corpora (Koester, 2022). Nonetheless, examining this link between reading and writing in a larger corpus of learner texts would be worth focusing on in further research. It would also be of interest to include a number of different topics, and perhaps different genres, in such a study, to examine any potential variation. In addition, such research might consider examining the extent to which general academic vocabulary present in the texts that learners have to read, is not produced. In other words, what proportion of the total amount of general academic vocabulary is being used by the learners.

Moreover, it would be interesting to investigate the existence of form similarity in general academic vocabulary, even further. In the present study, several examples of what Laufer (1988) defined as synforms were detected when evaluating correct use of general academic vocabulary. The fact that several of the words deemed incorrect were similar in form to the intended word, is worth considering. If this is examined in future research, and a larger corpus of learner texts is utilized, one might be able to detect a more systematic pattern of general academic words that can cause confusion among learners. Such words might be worthy of particular attention in the L2 English classroom.

Finally, it would be of interest to conduct further analyses of the relationship between the production of general academic vocabulary and achieved grades. Such research should consider gathering a larger sample in which inferential statistics may be used. Alternatively, it might utilize other aspects of descriptive statistics, such as confidence intervals or effect sizes. Moreover, research into this should strive to gather texts which have achieved grades across the entire grading scale, and overall, to avoid biases in the text selection process (Brezina, 2018).

As a final point, I would urge any further research into the topic, to be transparent about each step in the research process. This includes everything from specific software tools utilized, to what challenges that occurred, and how they were solved.



## References

- Adolphs, S., & Lin, P. M. (2011). Corpus linguistics. In J. Simpson, *The Routledge handbook of applied linguistics* (pp. 597-610). London: Routledge.
- Anthony, L. (2020). Resources for researching vocabulary. In S. Webb, *The Routledge handbook of vocabulary studies* (pp. 561-590). Oxon: Routledge.
- Anthony, L. (2022a). *TagAnt (Version 2.0.5) [Computer Software]*. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>.
- Anthony, L. (2022b). What can corpus software do? In A. O'Keeffe, & M. J. McCarthy, *The Routledge handbook of corpus linguistics (2nd ed.)* (pp. 103-125). London: Routledge.
- Anthony, L. (2023). *AntWordProfiler (Version 2.1.0) [Computer Software]*. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>.
- Ary, D., Jacobs, L. C., Sorensen, C. K., & Walker, D. A. (2014). *Introduction to research in education*. Wadsworth: Cengage Learning.
- Barclay, S., & Schmitt, N. (2019). Current perspectives on vocabulary teaching and learning. In X. Gao, *Second handbook of English language teaching* (pp. 1-22). Cham: Springer. [https://doi.org/10.1007/978-3-319-58542-0\\_42-1](https://doi.org/10.1007/978-3-319-58542-0_42-1).
- Bauer, L., & Nation, P. (1993). Word families. *International Journal of Lexicography*, 6(4), pp. 253-279.
- Baumann, J. F., & Graves, M. F. (2010). What is academic vocabulary? *Journal of Adolescent & Adult Literacy*, 54(1), pp. 4–12. doi:10.1598/JAAL.54.1.1.
- Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins Publishing Company.
- Biber, D. (2010). Corpus-based and corpus-driven analyses of language variation and use. In B. Heine, & H. Narrog, *The Oxford handbook of linguistic analysis* (pp. 159–192). Oxford: Oxford University Press.
- Biber, D., & Reppen, R. (2015). Introduction. In D. Biber, & R. Reppen, *The Cambridge handbook of English corpus linguistics* (pp. 1-8). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139764377.001>.
- Bonnet, G. (2004). *The assessment of pupils skills in English in eight European countries 2002*. Helsinki: The European network of policy makers for the evaluation of educational systems.
- Brevik, L. M., & Mathé, N. E. (2021). Mixed methods som forskningsdesign. In E. Andersson-Bakken, & C. P. Dalland, *Metoder i klasseromsforskning* (pp. 47-70). Oslo: Universitetsforlaget.
- Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781316410899>.
- Brezina, V., & McEnery, T. (2021). Introduction to corpus linguistics. In N. Tracy-Ventura, & M. Paquot, *The Routledge handbook of second language acquisition and corpora* (pp. 11-22). New York: Routledge.

- Busby, N. (2015). *Too cool for school? Sources of English language acquisition, attitudes, and academic reading ability among Norwegian university students*. [Master's thesis, Norwegian University of Science and Technology, Trondheim].
- Bybee, J. L., & Beckner, C. (2015). Usage-based theory. In B. Heine, & H. Narrog, *The Oxford handbook of linguistic analysis* (pp. 953-979). Oxford University Press: Oxford.
- Carlsen, C. (2012). Proficiency level—a fuzzy variable in computer learner corpora. *Applied Linguistics*, 33(2), pp. 161–183. doi:10.1093/applin/amr047.
- Carter, R. (1987). *Vocabulary: Applied linguistic perspectives*. London: Allen & Unwin.
- Celce-Murcia, M., & Schmitt, N. (2020). An overview of applied linguistics. In N. Schmitt, & M. P. Rodgers, *An introduction to applied linguistics* (pp. 1-15). London: Routledge.
- Chung, T. M., & Nation, P. (2003). Technical vocabulary in specialised texts. *Reading in a Foreign Language*, 15(2), pp. 103-116.
- Clark, V. L., & Ivankova, N. V. (2016). *Mixed methods research: A guide to the field*. Los Angeles: Sage.
- Cobb, T. (2010). Learning about language and learners from computer programs. *Reading in a Foreign Language*, 22(1), pp. 181–200.
- Cobb, T., & Horst, M. (2015). Learner corpora and lexis. In S. Granger, G. Gilquin, & F. Meunier, *The Cambridge handbook of learner corpus research* (pp. 185-206). Cambridge: Cambridge University Press.  
<https://doi.org/10.1017/CBO9781139649414.009>.
- Cons, A. M. (2012). The use and misuse of academic words in writing: Analyzing the writing of secondary English learners and redesignated learners. *TESOL Journal*, 3(4), pp. 610–638. doi: 10.1002/tesj.36.
- Corson, D. (1995). *Using English words*. Dordrecht: Springer Netherlands.
- Corson, D. (1997). The learning and use of academic English words. *Language Learning*, 47(4), pp. 671–718.
- Coxhead, A. (2000). A new Academic Word List. *TESOL Quarterly*, 34(2), pp. 213–238.  
<https://doi.org/10.2307/3587951>.
- Coxhead, A. (2011). The Academic Word List 10 years on: Research and teaching implications. *TESOL Quarterly*, 45(2), pp. 355 - 362.
- Coxhead, A. (2012). Academic vocabulary, writing and English for academic purposes: Perspectives from second language learners. *RELC Journal*, 43(1), pp. 137-145. DOI: 10.1177/0033688212439323.
- Coxhead, A. (2016a). Acquiring academic and disciplinary vocabulary. In K. Hyland, & P. Shaw, *The Routledge handbook of English for academic purposes* (pp. 177-190). London: Routledge.
- Coxhead, A. (2016b). Reflecting on Coxhead (2000), "A New Academic Word List". *TESOL Quarterly*, 50(1), pp. 181-185.
- Coxhead, A. (2018). *Vocabulary and English for specific purposes research: Quantitative and qualitative perspectives*. Oxon: Routledge.
- Coxhead, A. (2020). Academic vocabulary. In S. Webb, *The Routledge handbook of vocabulary studies* (pp. 97-110). Oxon: Routledge.

- Coxhead, A., & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing, 16*(3), pp. 129–147.
- Creswell, J. W., & Clark, V. L. (2018). *Designing and conducting mixed methods research (international student edition)*. Los Angeles: Sage.
- Crossley, S., Kyle, K., & Salsbury, T. (2016). A usage-based investigation of L2 lexical acquisition: The role of input and output. *The Modern Language Journal, 100*(3), pp. 702-715. <https://doi.org/10.1111/modl.12344>.
- Crossley, S., Salsbury, T., Titak, A., & McNamara, D. (2014). Frequency effects and second language lexical acquisition: Word types, word tokens, and word production. *International Journal of Corpus Linguistics, 19*(3), pp. 301–332. doi 10.1075/ijcl.19.3.01cro.
- Csomay, E. (2020). A corpus-based study of academic word use in EFL student writing. In U. Römer, V. Cortes, & E. Friginal, *Advances in corpus-based research on academic writing: Effects of discipline, register, and writer expertise* (pp. 9-32). Amsterdam: John Benjamins Publishing Company.
- Csomay, E., & Prades, A. (2018). Academic vocabulary in ESL student papers: A corpus-based study. *Journal of English for Academic Purposes, 33*, pp. 100-118.
- Dang, T. N. (2020). Corpus-based word lists in second language vocabulary research, learning, and teaching. In S. Webb, *The Routledge handbook of vocabulary studies* (pp. 288-303). Oxon: Routledge.
- Dang, T. N., Coxhead, A., & Webb, S. (2017). The Academic Spoken Word List. *Language Learning, 67*(4), pp. 959-997.
- Dang, T. N., Webb, S., & Coxhead, A. (2022). Evaluating lists of high-frequency words: Teachers' and learners' perspectives. *Language Teaching Research, 26*(4), pp. 617–641.
- DeKeyser, R. (2017). Knowledge and skill in ISLA. In S. Loewen, & M. Sato, *The Routledge handbook of instructed second language acquisition* (pp. 15-32). New York: Routledge.
- Dixon, T., Egbert, J., Larsson, T., Kaatari, H., & Hanks, E. (2023). Toward an empirical understanding of formality: Triangulating corpus data with teacher perceptions. *English for Specific Purposes, 71*, pp. 161-177. <https://doi.org/10.1016/j.esp.2023.04.006>.
- Durrant, P. (2009). Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes, 28*(3), pp. 157–169. doi:10.1016/j.esp.2009.02.002.
- Durrant, P. (2016). To what extent is the Academic Vocabulary List relevant to university student writing? *English for Specific Purposes, 43*, pp. 49-61.
- Durrant, P., Siyanova-Chanturia, A., Kremmel, B., & Sonbul, S. (2022). *Research methods in vocabulary studies*. Amsterdam: John Benjamins Publishing Company.
- Education First. (2023). *EF English proficiency index*. EF Education First. <https://www.ef.com/assetscdn/WIBIwq6RdJvcD9bc8RMd/cefcom-epi-site/reports/2023/ef-epi-2023-english.pdf>.

- Eldridge, J. (2008). "No, There Isn't an 'Academic Vocabulary,' but. . ." : A reader responds to K. Hyland and P. Tse's "Is There an 'Academic Vocabulary?'". *TESOL Quarterly*, 42(1), pp. 109-113.
- Ellis, N. (2018). Salience in usage-based SLA. In S. M. Gass, P. Spinner, & J. Behney, *Salience in second language acquisition* (pp. 21-40). New York: Routledge.
- Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), pp. 143–188. <https://doi.org/10.1017/S0272263102002024>.
- Ellis, N. C. (2012). Frequency-based accounts of second language acquisition. In S. M. Gass, & A. Mackey, *The Routledge handbook of second language acquisition* (pp. 193-210). London: Routledge.
- Ellis, N. C. (2013). Frequency effects. In P. Robinson, *The Routledge encyclopedia of second language acquisition* (pp. 260-265). New York: Routledge.
- Ellis, N. C. (2015). Cognitive and social aspects of learning from usage. In T. Cadierno, & S. W. Eskildsen, *Usage-based perspectives on second language learning* (pp. 49-73). Berlin: De Gruyter.
- Ellis, N. C. (2017). Cognition, corpora, and computing: Triangulating research in usage-based language learning. *Language Learning*, 67(S1), pp. 40-65. DOI: 10.1111/lang.12215.
- Ellis, N. C., & Wulff, S. (2020). Usage-based approaches to L2 acquisition. In B. VanPatten, G. D. Keating, & S. Wulff, *Theories in second language acquisition: An introduction* (pp. 63-82). New York: Routledge.
- Ellis, N. C., Römer, U., & O'Donnell, M. B. (2016). *Usage-based approaches to language acquisition and processing: Cognitive and corpus investigations of construction grammar*. Malden, MA: Wiley.
- Ellis, N., & Beaton, A. (1993). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *The Quarterly Journal of Experimental Psychology Section A*, 46(3), pp. 533-558. DOI: 10.1080/14640749308401062.
- Ellis, R. (2008). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2009). Implicit and explicit learning, knowledge and instruction. In R. Ellis, S. Loewen, C. Elder, H. Reinders, R. Erlam, & J. Philp, *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 3-25). Buffalo: Multilingual Matters.
- Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. Oxon: Routledge.
- Flowerdew, L. (2004). The argument for using English specialized corpora to understand academic and professional language. In U. Connor, & T. A. Upton, *Discourse in the professions: Perspectives from corpus linguistics* (pp. 11-33). Amsterdam: John Benjamins Publishing Company.
- Flowerdew, L. (2012). *Corpora and language education*. London: Palgrave Macmillan.
- Flowerdew, L. (2015). Learner corpora and language for academic and specific purposes. In S. Granger, G. Gilquin, & F. Meunier, *The Cambridge handbook of learner corpus research* (pp. 465-484). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139649414.021>.

- Gablasova, D., & Brezina, V. (2021). Words that matter in L2 research and pedagogy: A corpus-linguistics perspective. *Studies in Second Language Acquisition*, 43, pp. 958–961. doi:10.1017/S027226312100070X.
- Gardner, D. (2007). Validating the construct of word in applied corpus-based vocabulary research: A critical survey. *Applied Linguistics*, 28(2), pp. 241–265. doi:10.1093/applin/amm010.
- Gardner, D., & Davies, M. (2014). A new Academic Vocabulary List. *Applied Linguistics*, 35(3), pp. 305–327. doi:10.1093/applin/amt015.
- Gardner, D., & Davies, M. (2016). A response to “To what extent is the Academic Vocabulary List relevant to university student writing?”. *English for Specific Purposes*, 43, pp. 62–68. <https://doi.org/10.1016/j.esp.2016.03.001>.
- Gass, S. M., Spinner, P., & Behney, J. (2018). Salience in second language acquisition and related fields. In S. M. Gass, P. Spinner, & J. Behney, *Salience in second language acquisition* (pp. 1–18). New York: Routledge.
- Gilner, L. (2011). A primer on the General Service List. *Reading in a Foreign Language*, 23(1), pp. 65–83.
- Gilquin, G., Granger, S., & Paquot, M. (2007). Learner corpora: The missing link in EAP pedagogy. *Journal of English for Academic Purposes*, 6(4), pp. 319–335. <https://doi.org/10.1016/j.jeap.2007.09.007>.
- Granger, S. (2002). A bird’s-eye view of learner corpus research. In S. Granger, J. Hung, & S. Petch-Tyson, *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 3–33). Amsterdam: John Benjamins Publishing Company.
- Granger, S. (2009). The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation. In K. Aijmer, *Corpora and language teaching* (pp. 13–32). Amsterdam: John Benjamins Publishing Company.
- Granger, S., & Larsson, T. (2021). Is core vocabulary a friend or foe of academic writing? Single-word vs multi-word uses of THING. *Journal of English for Academic Purposes*, 52, pp. 1–13. <https://doi.org/10.1016/j.jeap.2021.100999>.
- Gries, S. T. (2008). Corpus-based methods in analyses of second language acquisition data. In P. Robinson, & N. C. Ellis, *Handbook of cognitive linguistics and second language acquisition* (pp. 406–431). New York: Routledge.
- Gries, S. T. (2015). Statistics for learner corpus research. In S. Granger, G. Gilquin, & F. Meunier, *The Cambridge handbook of learner corpus research* (pp. 159–182). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139649414.008>.
- Gries, S. T. (2022). How to use statistics in quantitative corpus analysis. In A. O’Keeffe, & M. J. McCarthy, *The Routledge handbook of corpus linguistics* (pp. 168–181). London: Routledge.
- 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. *The Journal of Language Teaching and Learning*, 6(2), pp. 70–87.
- Hasselgren, A. (1994). Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary. *International Journal of Applied Linguistics*, 4(2), pp. 237–258.

- 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?* [Doktorgradsavhandling]. Universitetet i Oslo.
- Hellekjær, G. O. (2008). A case for improved reading instruction for academic English reading proficiency. *Acta Didactica Norge*, 2(1), pp. 1-17.
- Hellekjær, G. O. (2009). Academic English reading proficiency at the university level: A Norwegian case study. *Reading in a Foreign Language*, 21(2), pp. 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, *Lexical issues in L2 writing* (pp. 29-56). Newcastle upon Tyne: Cambridge Scholars Publishing.
- Henriksen, B., Holmen, A., & Kling, J. (2019). *English medium instruction in multilingual and multicultural universities: Academics' voices from the Northern European context*. London: Routledge.
- Hestetræet, T. I. (2020). Learning and teaching vocabulary. In A.-B. Fenner, & A. S. Skulstad, *Teaching English in the 21st century: Central issues in English didactics* (pp. 185-211). Bergen: Fagbokforlaget.
- Hinkel, E. (2003). Simplicity without elegance: Features of sentences in L1 and L2 academic texts. *TESOL Quarterly*, 37(2), pp. 275-301. <https://doi.org/10.2307/3588505>.
- Hirsh, D. (2010). Researching vocabulary. In B. Paltridge, & A. Phakiti, *Continuum companion to research methods in applied linguistics* (pp. 222-255). London: Continuum.
- Hirvela, A. (2016). Academic reading into writing. In K. Hyland, & P. Shaw, *The Routledge handbook of English for academic purposes* (pp. 127-138). London: Routledge.
- Hudson, T. (2015). Presenting quantitative data visually. In L. Plonsky, *Advancing quantitative methods in second language research* (pp. 78-105). New York: Routledge.
- Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson, *Cognition and second language instruction* (pp. 258-286). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524780.011>.
- Hulstijn, J. H., Alderson, J. C., & Schoonen, R. (2010). Developmental stages in second-language acquisition and levels of second-language proficiency: Are there links between them? In I. Bartning, M. Martin, & I. Vedder, *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (pp. 11-20). European Second Language Association: EUROSLA Monographs Series.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Hyland, K. (2010). Researching writing. In B. Paltridge, & A. Phakiti, *Continuum companion to research methods in applied linguistics* (pp. 191-204). London: Continuum.
- Hyland, K., & Tse, P. (2007). Is There an “Academic Vocabulary”? *TESOL Quarterly*, 41(2), pp. 235-253. <https://doi.org/10.1002/j.1545-7249.2007.tb00058.x>.
- Johnson, R. B., & Christensen, L. (2020). *Educational research: Quantitative, qualitative, and mixed approaches*. Los Angeles: SAGE.

- Kamaşak, R., Sahan, K., & Rose, H. (2021). Academic language-related challenges at an English-medium university. *Journal of English for Academic Purposes*, 49. <https://doi.org/10.1016/j.jeap.2020.100945>.
- Kemmer, S., & Barlow, M. (2000). Introduction: A usage-based conception of language. In M. Barlow, & S. Kemmer, *Usage-based models of language* (pp. vii-xxviii). Stanford: CSLI Publications.
- Koester, A. (2022). Building small specialised corpora. In A. O'Keeffe, & M. J. McCarthy, *The Routledge handbook of corpus linguistics (2nd ed.)* (pp. 48-61). London: Routledge.
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, 73(4), pp. 440-464.
- Kyle, K. (2021). Lexis. In N. Tracy-Ventura, & M. Paquot, *The Routledge handbook of second language acquisition and corpora* (pp. 332-344). New York: Routledge.
- Landsberger, H. A. (1958). *Hawthorne revisited: Management and the worker, its critics, and developments in human relations in industry*. Ithaca, New York: Cornell University.
- Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow, & S. Kemmer, *Usage-based models of language* (pp. 1-63). Standford: CSLI Publications.
- Langacker, R. W. (2008). Cognitive grammar as a basis for language instruction. In P. Robinson, & N. C. Ellis, *Handbook of cognitive linguistics and second language acquisition* (pp. 66–88). London: Routledge.
- Langacker, R. W. (2009). A dynamic view of usage and language acquisition. *Cognitive Linguistics*, 20(3), pp. 627–640. DOI 10.1515/COGL.2009.027.
- Larson-Hall, J., & Herrington, R. (2009). Improving data analysis in second language acquisition by utilizing modern developments in applied statistics. *Applied Linguistics*, 31(3), pp. 368–390. <https://doi.org/10.1093/applin/amp038>.
- Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning*, 65(1), pp. 127-159. DOI: 10.1111/lang.12115.
- Larsson, T., Paquot, M., & Plonsky, L. (2020). Inter-rater reliability in learner corpus research: Insights from a collaborative study on adverb placement. *International Journal of Learner Corpus Research*, 6(2), pp. 237-251. <https://doi.org/10.1075/ijlcr.20001.lar>.
- Laufer, B. (1988). The concept of “synforms” (similar lexical forms) in L2 vocabulary acquisition. *Language and Education*, 2(2), pp. 113-132.
- Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics*, 19(2), pp. 255–271. <https://doi.org/10.1093/applin/19.2.255>.
- Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA Yearbook*, 5(1), pp. 223-250.
- Laufer, B. (2013). Vocabulary and writing. In Chapelle, & C. A., *The encyclopedia of applied linguistics*. Blackwell Publishing Ltd. DOI: 10.1002/9781405198431.wbeal1432.
- Laufer, B. (2021). Lemmas, flemmas, word families and common sense. *Studies in Second Language Acquisition*, 43, pp. 965–968. doi:10.1017/S0272263121000656 .

- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), pp. 307-322.
- Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, 48(3), pp. 365-391.
- Lee, S. H. (2003). ESL learners' vocabulary use in writing and the effects of explicit vocabulary instruction. *System*, 31, pp. 537-561.
- Leki, I., & Carson, J. (1997). "Completely different worlds": EAP and the writing experiences of ESL students in university courses. *TESOL Quarterly*, 31(1), pp. 39-69. <https://doi.org/10.2307/3587974>.
- Lillis, T., & Scott, M. (2007). Defining academic literacies research: Issues of epistemology, ideology and strategy. *Journal of Applied Linguistics*, 4(1), pp. 5-32. <https://doi.org/10.1558/japl.v4i1.5>.
- Lin, L. H., & Morrison, B. (2010). The impact of the medium of instruction in Hong Kong secondary schools on tertiary students' vocabulary. *Journal of English for Academic Purposes*, 9(4), pp. 255-266. <https://doi.org/10.1016/j.jeap.2010.09.002>.
- Ljosland, R. (2007). English in Norwegian academia - a step towards diglossia? *World Englishes*, 26(4), pp. 395-410. <https://doi.org/10.1111/j.1467-971X.2007.00519.x>.
- Loewen, S., & Plonsky, L. (2016). *An A - Z of applied linguistics research methods*. London: Palgrave.
- Loewen, S., & Reinders, H. (2011). *Key concepts in second language acquisition*. Hampshire, UK: Palgrave Macmillan.
- Mahan, K. R. (2013). *Lexical errors in Norwegian intermediate and advanced learners of English*. [Masteroppgave]. Universitetet i Oslo.
- Malmström, H., Pecorari, D., & Gustafsson, M. (2016). Coverage and development of academic vocabulary in assessment texts in English medium instruction. In S. Göpferich, & I. Neumann, *Developing and assessing academic and professional writing skills* (pp. 45-69). Frankfurt am Main: Peter Lang International Academic Publishing Group.
- Malmström, H., Pecorari, D., & Shaw, P. (2018). Words for what? Contrasting university students' receptive and productive academic vocabulary needs. *English for Specific Purposes*, 50, pp. 28-39. <https://doi.org/10.1016/j.esp.2017.11.002>.
- Martin, M. (1984). Advanced vocabulary teaching: The problem of synonyms. *The Modern Language Journal*, 68(2), pp. 130-137. <https://doi.org/10.2307/327139>.
- Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics*, 33(3), pp. 299-320. doi:10.1093/applin/ams010.
- Martinez, R., & Schmitt, N. (2015). Vocabulary. In D. Biber, & R. Reppen, *The Cambridge handbook of English corpus linguistics* (pp. 439-459). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139764377.025>.
- 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, pp. 44-57. <https://doi.org/10.1016/j.jeap.2017.12.006>.



- Mauranen, A., Hynninen, N., & Ranta, E. (2010). English as an academic lingua franca: The ELFA project. *English for Specific Purposes*, 29(3), pp. 183-190.  
<https://doi.org/10.1016/j.esp.2009.10.001>.
- Mauranen, A., Hynninen, N., & Ranta, E. (2016). English as the academic lingua franca. In K. Hyland, & P. Shaw, *The Routledge handbook of English for academic purposes* (pp. 44-55). London: Routledge.
- McEnery, T., & Brookes, G. (2022). Building a written corpus: What are the basics? In A. O'Keeffe, & M. J. McCarthy, *The Routledge handbook of corpus linguistics (2nd ed.)* (pp. 35-47). London: Routledge.
- McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge: Cambridge University Press.
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), pp. 276-282. <https://doi.org/10.11613/BM.2012.031>.
- Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching and Linguistics: Abstracts*, 13(4), pp. 221–246.
- Milton, J. (2001). Elements of a written interlanguage: A computational and corpus-based study of institutional influences on the acquisition of English by Hong Kong Chinese students. Research Reports (Vol. 2). [https:// repository.ust.hk/ir/bitstream/1783.1-1055/1/john.pdf](https://repository.ust.hk/ir/bitstream/1783.1-1055/1/john.pdf).
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol: Multilingual Matters.
- Milton, J., Wade, J., & Hopkins, N. (2010). Aural word recognition and oral competence in English as a foreign language. In R. Chacón-Beltrán, C. Abello-Contesse, & M. d. Torreblanca-López, *Insights into non-native vocabulary teaching and learning* (pp. 83-98). Bristol: Multilingual Matters.
- Ministry of Education and Research. (2019). *Curriculum in English (ENG01-04)*. Established as a regulation. National curriculum for Knowledge Promotion 2020. Retrieved from <https://www.udir.no/lk20/eng01-04>.
- Mitchell, R., Myles, F., & Marsden, E. (2019). *Second language learning theories*. New York: Routledge.
- Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, 47(1), pp. 91–108.  
<https://doi.org/10.1002/RRQ.011>.
- Nation, P. (1983). Testing and teaching vocabulary. *Guidelines*, 5(1), pp. 12-25.
- Nation, P. (2011). Research into practice: Vocabulary. *Language Teaching*, 44(4), pp. 529 - 539.
- Nation, P. (2016). *Making and using word lists for language learning and testing*. Amsterdam: John Benjamins Publishing Company.
- Nation, P. (2017). *The BNC/COCA level 6 word family lists (Version 1.0.0) [Data file]*. Retrieved from <http://www.victoria.ac.nz/lals/staff/paul-nation.aspx>
- Nation, P. (2022). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, P., & Meara, P. (2020). Vocabulary. In N. Schmitt, & M. P. Rodgers, *An introduction to applied linguistics* (pp. 35-54). London: Routledge.

- Nation, P., Coxhead, A., Chung, T. M., & Quero, B. (2016). Specialized word lists. In P. Nation, *Making and using word lists for language learning and testing* (pp. 145-151). Amsterdam: John Benjamins Publishing Company.
- Nesi, H. (2016). Corpus studies in EAP. In K. Hyland, & P. Shaw, *The Routledge handbook of English for academic purposes* (pp. 206-217). London: Routledge.
- Norwegian Directorate for Education and Training. (2023). *Fag- og trefordeling og tilbudsstruktur for Kunnskapsløftet*. (Udir-1-2023) [Rundskriv]. Retrieved from <https://www.udir.no/regelverkstolkninger/opplaring/Innhold-i-opplaringen/udir-1-2023/>
- Norwegian Universities and Colleges Admission Service. (2024, February 22). *Norwegian and English language requirements*. Retrieved from <https://www.samordnaopptak.no/info/english/language-requirements/>
- Oakey, D. (2022). What can a corpus tell us about lexis? In A. O'Keeffe, & M. J. McCarthy, *The Routledge handbook of corpus linguistics* (pp. 185-203). London: Routledge. DOI: 10.4324/9780367076399-14.
- O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge: Cambridge University Press.
- Olsson, E. (2015). Progress in English academic vocabulary use in writing among CLIL and non-CLIL students in Sweden. *Moderna Språk, 109*(2), pp. 51–74. <https://doi.org/10.58221/mosp.v109i2.7927>.
- Olsson, E., & Sylvé, 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), pp. 77–103. DOI: 10.17011/apples/urn.201512234129.
- Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Education.
- Ortega, L. (2011). Second language acquisition. In J. Simpson, *The Routledge handbook of applied linguistics* (pp. 171-184). London: Routledge.
- Ortega, L. (2015). Usage-based SLA: A research habitus whose time has come. In T. Cadierno, & S. W. Eskildsen, *Usage-based perspectives on second language learning* (pp. 353-373). Berlin: De Gruyter.
- Paquot, M. (2010). *Academic vocabulary in learner writing*. London: Continuum.
- Paquot, M., & Plonsky, L. (2017). Quantitative research methods and study quality in learner corpus research. *International Journal of Learner Corpus Research, 3*(1), pp. 61-94. DOI: 10.1075/ijlcr.3.1.03paq.
- Plakans, L., & Gebril, A. (2012). A close investigation into source use in integrated second language writing tasks. *Assessing Writing, 17*(1), pp. 18–34. <https://doi.org/10.1016/j.asw.2011.09.002>.
- Plonsky, L. (2013). Study quality in SLA: An assessment of designs, analyses, and reporting practices in quantitative L2 research. *Studies in Second Language Acquisition, 35*(4), pp. 655–687. doi:10.1017/S0272263113000399.
- Plonsky, L. (2014). Study quality in quantitative L2 research (1990–2010): A methodological synthesis and call for reform. *The Modern Language Journal, 98*(1), pp. 450-470. <https://doi.org/10.1111/j.1540-4781.2014.12058.x>.

- Qian, D. D., & Lin, L. H. (2020). The relationship between vocabulary knowledge and language proficiency. In S. Webb, *The Routledge handbook of vocabulary studies* (pp. 66-80). Oxon: Routledge.
- Römer, U. (2013). Corpus analysis. In P. Robinson, *The Routledge encyclopedia of second language acquisition* (pp. 134-139). New York: Routledge.
- Römer, U. (2023). Usage-based approaches to second language acquisition vis-à-vis data-driven learning. *TESOL Quarterly*, pp. 1-11. <https://doi.org/10.1002/tesq.3278>.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Reppen, R. (2022). Building a corpus: What are key considerations? In A. O'Keeffe, & M. J. McCarthy, *The Routledge handbook of corpus linguistics (2nd ed.)* (pp. 13-20). London: Routledge.
- Reppen, R., & Simpson-Vlach, R. (2020). Corpus linguistics. In N. Schmitt, & M. P.H. Rodgers, *An introduction to applied linguistics* (pp. 91-108). London: Routledge.
- Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly*, 10(1), pp. 77-89. <https://doi.org/10.2307/3585941>.
- Rindal, U., & Brevik, L. M. (2019). State of the art: English didactics in Norway. In U. Rindal, & L. M. Brevik, *English Didactics in Norway – 30 years of doctoral research* (pp. 418-440). Universitetsforlaget. DOI: 10.18261/978-82-15-03074-6-2019-21.
- Robinson, P., & Ellis, N. C. (2008). Conclusion: Cognitive linguistics, second language acquisition and L2 instruction—issues for research. In P. Robinson, & N. C. Ellis, *Handbook of cognitive linguistics and second language acquisition* (pp. 489-545). New York: Routledge.
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt, *Attention and awareness in foreign language learning* (pp. 1-63). Honolulu: University of Hawaii: Second Language Teaching and Curriculum Center.
- Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), pp. 129-158.
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), pp. 329–363. <https://doi.org/10.1177/1362168808089921>.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke: Palgrave Macmillan.
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows? *Language Learning*, 64(4), pp. 913-951. DOI: 10.1111/lang.12077.
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(4), pp. 484 - 503. <https://doi.org/10.1017/S0261444812000018>.
- Schmitt, N., & Schmitt, D. (2020). *Vocabulary in language teaching: Second edition*. Cambridge: Cambridge University Press.
- Schmitt, N., Cobb, T., Horst, M., & Schmitt, D. (2015). How much vocabulary is needed to use English? Replication of van Zeeland & Schmitt (2012), Nation (2006) and Cobb (2007). *Language Teaching*, 50(2), pp. 212–226. doi:10.1017/S0261444815000075.

- 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), pp. 55-88.
- Seidlhofer, B. (2003). *Controversies in applied linguistics*. Oxford: Oxford University Press.
- Siyanova-Chanturia, A., & Webb, S. (2016). Teaching vocabulary in the EFL context. In W. A. Renandya, & H. P. Widodo, *English language teaching today: Linking theory and practice* (pp. 227-239). Switzerland: Springer International Publishing.
- Skarpaas, K. G. (2011). *Subject choice and the English programme subjects in Norwegian upper secondary school: A survey*. [Master's thesis, University of Oslo, Norway]. <http://urn.nb.no/URN:NBN:no-29273>.
- Skjelde, K. M., & Coxhead, A. (2020). Mind the gap: Academic vocabulary knowledge as a predictor of English grades. *Acta Didactica Norden*, 14(3), pp. 1-20. <https://doi.org/10.5617/adno.7975>.
- Skulstad, A. S. (2020). Communicative competence. In A.-B. Fenner, & A. S. Skulstad, *Teaching English in the 21st century: Central issues in English didactics* (pp. 43-67). Bergen: Fagbokforlaget.
- Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson, & N. Torrance, *The Cambridge handbook of literacy* (pp. 112–133). New York: Cambridge University Press.
- Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, 36(2), pp. 139-152. DOI: 10.1080/09571730802389975.
- Stoeckel, T., Ishii, T., & Bennett, P. (2020). Is the lemma more appropriate than the flemma as a word counting unit? *Applied Linguistics*, 41(4), pp. 601–606. doi:10.1093/applin/amy059.
- Szudarski, P. (2018). *Corpus linguistics for vocabulary: A guide for research*. Oxon: Routledge.
- Tang, R. (2012). The issues and challenges facing academic writers from ESL/EFL contexts: An overview. In R. Tang, *Academic writing in a second or foreign language: Issues and challenges facing ESL/EFL academic writers in higher education contexts* (pp. 1-18). New York: Continuum.
- Therova, D. (2021). A case study of academic vocabulary in a novice student's writing at a UK university. *JELTL (Journal of English Language Teaching and Linguistics)*, 6(3), pp. 557-577. <https://dx.doi.org/10.21462/jeltl.v6i3.609>.
- Therova, D. (2022). The development of academic vocabulary in international foundation students' assessed academic writing. *Journal of International Students*, 12(1), pp. 123-140. doi: 10.32674/jis.v12i1.2824.
- Therova, D. (2023). Core academic vocabulary in four genres of novice student writing. *Teaching English as a Second Language Electronic Journal (TESL-EJ)*, 27(1), pp. 1-21. <https://doi.org/10.55593/ej.27105a8>.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, Massachusetts: Harvard University Press.
- Tyler, A. (2010). Usage-based approaches to language and their applications to second language learning. *Annual Review of Applied Linguistics*, 30, pp. 270–291. doi:10.1017/S0267190510000140.

- Tyler, A. E., & Ortega, L. (2018a). Usage-inspired L2 instruction: An emergent, researched pedagogy. In A. E. Tyler, L. Ortega, M. Uno, & H. I. Park, *Usage-inspired L2 instruction: Researched pedagogy* (pp. 3–26). Amsterdam: John Benjamins Publishing Company.
- Tyler, A. E., & Ortega, L. (2018b). Usage-inspired L2 instruction: Some reflections and a heuristic. In A. E. Tyler, L. Ortega, M. Uno, & H. I. Park, *Usage-inspired L2 instruction: Researched pedagogy* (pp. 315-321). Amsterdam: John Benjamins Publishing Company.
- VanPatten, B., & Benati, A. G. (2010). *Key terms in second language acquisition*. London: Continuum.
- Verspoor, M., & Schmitt, N. (2013). Language and the lexicon in SLA. In P. Robinson, *The Routledge encyclopedia of second language acquisition* (pp. 353-360). New York: Routledge.
- Verspoor, M., & Smiskova, H. (2012). Foreign language writing development from a dynamic usage based perspective. In R. M. Manchón, *L2 writing development: Multiple perspectives* (pp. 17-46). Boston: De Gruyter Mouton.
- Vilkaitė-Lozdienė, L., & Schmitt, N. (2020). Frequency as a guide for vocabulary usefulness. In S. Webb, *The Routledge handbook of vocabulary studies* (pp. 81-96). Oxon: Routledge.
- Warnby, M. (2023). Academic vocabulary knowledge among adolescents in university preparatory programmes. *Journal of English for Academic Purposes, 61*. <https://doi.org/10.1016/j.jeap.2022.101203>.
- Webb, S. (2008). RECEPTIVE AND PRODUCTIVE VOCABULARY SIZES OF L2 LEARNERS. *Studies in Second Language Acquisition, 30*(1), pp. 79-95.
- Webb, S. (2021). Word families and lemmas, not a real dilemma: Investigating lexical units. *Studies in Second Language Acquisition, 43*, pp. 973–984. doi:10.1017/S0272263121000760.
- Webb, S., & Nation, P. (2017). *How vocabulary is learned?* Oxford: Oxford University Press.
- 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), pp. 33–69.
- West, M. (1953). *A general service list of English words*. London: Longmans, Green and co.
- Wulff, S. (2021). Usage-based approaches. In N. Tracy-Ventura, & M. Paquot, *The Routledge handbook of second language acquisition and corpora* (pp. 175-188). New York: Routledge.
- Wulff, S., & Ellis, N. C. (2018). Usage-based approaches to SLA. In D. Miller, F. Bayram, J. Rothman, & L. Serratrice, *Bilingual cognition and language: The state of the science across its subfields* (pp. 37-56). Amsterdam: John Benjamins.
- Xue, G., & Nation, P. (1984). A University Word List. *Language Learning and Communication, 3*(2), pp. 215-229.
- Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. In J. Coady, & T. Huckin, *Second language vocabulary acquisition* (pp. 5-19). Cambridge: Cambridge University Press.

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. New York: Hafner Publishing Company.

## Appendices

### Appendix A

Letter of information and consent form

## **Are you interested in taking part in the research project: “Use of Academic Vocabulary in Written English Mock Exams”?**

### **Purpose of the project**

You are invited to participate in a research project where the purpose is to look at:

- how many academic words are used in written English mock exam texts by learners in vg1,
- whether these academic words are used correctly, and
- the relationship between the use of academic words and the written English mock exam grade.

The reason why it is interesting to look into learners' use of academic words, is because knowing how to understand and use this vocabulary is important in order to be successful at university, no matter what study you aim at. This research project is a master's thesis written in English didactics, as part of the teacher education programme at the University of Bergen.

### **Which institution is responsible for the research project?**

The University of Bergen is responsible for the project (data controller).

### **Why are you being asked to participate?**

The reason why you have been asked to participate is because you have written a text in the English mock exam during vg1 last spring, in a Norwegian upper secondary school at the general studies programme. This is the only criterion for being asked to participate in this research project.

### **What does participation involve for you?**

This research project is a corpus study, which means that your written mock exam text will be part of an electronic collection of texts which will be analysed. Participating in this research project involves allowing your teacher to provide me with your written English mock exam text from vg1, and the grade you achieved on this mock exam.

**Participation is voluntary**

Participation in the project is voluntary. If you choose to participate, you can withdraw your consent at any time without giving a reason. All information about you will then be made anonymous. There will be no negative consequences for you if you chose not to participate or later decide to withdraw.

**Your personal privacy – how we will store and use your personal data**

We will only use your personal data for the purpose(s) specified here and we will process your personal data in accordance with data protection legislation (the GDPR).

- Master student Filip Thorvaldsen and supervisor Kimberly Skjelde will have access to the personal data collected in this research project.
- Your name will not be used at all in this research project. Information about which learner has written which text and received which grade, will be replaced with a code.
- Participants will not be recognizable in the publication of this research project. The only type of information that will be published will be information about academic words used in your texts, in addition to the relationship between this and grades. At most, single sentences may be published in order to provide context for the use of specific academic words.

**What will happen to your personal data at the end of the research project?**

The planned end date of the project is the 15<sup>th</sup> of May 2024. When the project ends, your written English mock exam text and information about your grade will be deleted.

**Your rights**

So long as you can be identified in the collected data, you have the right to:

- access the personal data that is being processed about you
- request that your personal data is deleted
- request that incorrect personal data about you is corrected/rectified
- receive a copy of your personal data (data portability), and
- send a complaint to the Norwegian Data Protection Authority regarding the processing of your personal data

**What gives us the right to process your personal data?**

We will process your personal data based on your consent.

**Where can I find out more?**

If you have questions about the project, or want to exercise your rights, contact:

- University of Bergen via Filip Thorvaldsen (student) at [filip.thorvaldsen@student.uib.no](mailto:filip.thorvaldsen@student.uib.no) or Kimberly Skjelde (supervisor) at [Kimberly.skjelde@uib.no](mailto:Kimberly.skjelde@uib.no)



Yours sincerely,

Kimberly Skjelde  
Project Leader (supervisor)

Filip Thorvaldsen  
Student

---

## Consent form

I have received and understood information about the project “Use of Academic Vocabulary in Written English Mock Exams” and have been given the opportunity to ask questions. I give consent:

- to participate in this research project
- for my teacher to give information about me to this project

**In order to get your written English mock exam text from vg1, and the grade you got on this mock exam, state the name of your English teacher at the time:**

The name of my English teacher in vg1 is:

---

I give consent for my personal data to be processed until the end of the project.

---

---

(Signed by participant, date)

## Appendix B

### Tokens in the Corpus

<b>Learner ID</b>	<b>Task 1</b>	<b>Task 2</b>	<b>Task 3</b>	<b>Total</b>
<i>AA</i>	208	265	509	982
<i>AB</i>	201	259	555	1,015
<i>AC</i>	189	257	614	1,060
<i>AD</i>	197	252	690	1,139
<i>AE</i>	156	249	692	1,097
<i>AF</i>	207	248	732	1,187
<i>AG</i>	199	234	693	1,126
<i>AH</i>	173	276	548	997
<i>AI</i>	201	266	650	1,117
<i>AJ</i>	254	219	608	1,081
<i>AK</i>	193	249	736	1,178
<i>AL</i>	261	156	676	1,093
<i>AM</i>	170	255	654	1,079
<i>AN</i>	206	244	618	1,068
<i>AO</i>	241	260	786	1,287
<i>AP</i>	193	210	613	1,016
<i>AQ</i>	163	266	802	1,231
<i>AR</i>	195	233	626	1,054
<i>AS</i>	222	311	589	1,122
<i>AT</i>	213	298	816	1,327
<i>AU</i>	199	255	723	1,177
<i>AV</i>	162	246	412	820
<i>BA</i>	188	189	445	822
<i>BB</i>	201	210	1,031	1,442
<i>BC</i>	200	193	926	1,319
<i>BD</i>	190	201	832	1,223
<i>BE</i>	200	181	777	1,158
<i>BF</i>	195	198	499	892
<b>Total</b>	5,577	6,680	18,852	31,109

*Note.* The right column is the total for each learner (i.e. task 1-3). The bottom row is the total for all texts belonging to each task.

**Appendix C**  
Types in the Corpus

<b>Learner ID</b>	<b>Task 1</b>	<b>Task 2</b>	<b>Task 3</b>	<b>Total</b>
<i>AA</i>	120	135	196	338
<i>AB</i>	112	138	242	408
<i>AC</i>	113	136	240	378
<i>AD</i>	127	121	249	377
<i>AE</i>	100	146	251	402
<i>AF</i>	108	129	240	376
<i>AG</i>	95	116	215	325
<i>AH</i>	101	144	217	361
<i>AI</i>	111	143	270	422
<i>AJ</i>	131	116	237	384
<i>AK</i>	114	124	308	468
<i>AL</i>	121	89	259	375
<i>AM</i>	99	132	272	395
<i>AN</i>	125	124	285	445
<i>AO</i>	125	129	298	430
<i>AP</i>	96	123	203	337
<i>AQ</i>	113	154	299	470
<i>AR</i>	96	113	218	336
<i>AS</i>	126	140	220	373
<i>AT</i>	110	123	256	373
<i>AU</i>	112	145	290	439
<i>AV</i>	104	138	185	329
<i>BA</i>	100	117	196	312
<i>BB</i>	104	114	345	449
<i>BC</i>	109	112	329	444
<i>BD</i>	97	125	288	399
<i>BE</i>	104	99	262	335
<i>BF</i>	97	116	178	301
<b>Total</b>	866	1,099	2,362	3,033

### Appendix D

Hyphenated Words on the AVL (Gardner & Davies, 2014) and their Frequency Number on  
the List

373	Long-term (j)	2043	Top-down (j)	2623	State-level (j)
719	Well-being (n)	2091	Computer-based (j)	2657	Agreed-upon (j)
891	Short-term (j)	2157	Self-sufficiency (n)	2699	Peer-reviewed (j)
999	Large-scale (j)	2163	Re-evaluate (v)	2705	Cross-national (j)
1207	Long-standing (j)	2168	Post-traumatic (j)	2709	De-emphasize (v)
1232	Decision-making (j)	2186	Broad-based (j)	2711	Self-sacrifice (n)
1256	Decision-making (n)	2205	Self-evident (j)	2718	Self-help (j)
1328	High-quality (j)	2284	Long-lasting (j)	2729	Add-on (j)
1372	In-depth (j)	2286	Labor-intensive (j)	2730	Near-term (j)
1487	Trade-off (n)	2309	Non-white (n)	2743	Well-organized (j)
1498	Face-to-face (j)	2310	Longer-term (j)	2756	Non-competitive (j)
1527	Present-day (j)	2311	Self-awareness (n)	2769	City-state (n)
1566	On-site (j)	2334	Cross-section (n)	2776	Second-order (j)
1570	By-product (n)	2345	Co-author (v)	2796	Male-female (j)
1576	Cost-effective (j)	2349	Self-serving (j)	2811	Gender-specific (j)
1606	Self-interest (n)	2365	Two-fold (j)	2818	All-encompassing (j)
1610	High-level (j)	2374	Ever-increasing (j)	2826	Self-defeating (j)
1665	Two-way (j)	2375	Self-reliance (n)	2834	First-order (j)
1686	Real-world (j)	2403	Well-developed (j)	2853	Add-on (n)
1689	One-way (j)	2424	Ten-year (j)	2882	Two-stage (j)
1730	Small-scale (j)	2426	Re-evaluation (n)	2891	Two-step (j)
1791	Far-reaching (j)	2428	Well-documented (j)	2897	Self-preservation (n)
1815	In-house (j)	2456	Value-added (j)	2900	Long-established (j)
1824	Third-party (j)	2488	Second-generation (j)	2927	Self-selected (j)
1831	Re-establish (v)	2505	One-dimensional (j)	2935	Fine-grained (j)
1834	Time-consuming (j)	2523	Zero-sum (j)	2937	Military-industrial (j)
1857	Re-examine (v)	2524	Upper-class (j)	2938	Post-industrial (j)
1917	English-speaking (j)	2550	Re-examination (n)	2952	Cause-and-effect (j)
1944	Self-image (n)	2553	Deep-seated (j)	2961	Market-driven (j)
1946	Decision-maker (n)	2555	Well-designed (j)	2971	Self-improvement (n)
1972	Pre-existing (j)	2557	Above-mentioned (j)	2983	Upper-level (j)
1987	Well-established (j)	2598	Laissez-faire (n)	2985	High-value (j)
2036	Well-defined (j)	2615	State-sponsored (j)	2995	Non-professional (j)

Note. n = noun, v = verb, j = adjective, adv = adverb.

### Appendix E

Words Appearing on the AVL (Gardner & Davies, 2014) More than Once

1	Aggregate	26	Sequence	51	Component	76	Increase
2	Subordinate	27	Review	52	Characteristic	77	Female
3	Parallel	28	Radical	53	Bias	78	Equal
4	Decision-making	29	Prior	54	Author	79	Effect
5	Add-on	30	Potential	55	Attribute	80	Design
6	Outreach	31	Overall	56	Approximate	81	Conditional
7	Conjecture	32	Objective	57	Appropriate	82	Condition
8	Postulate	33	Multiple	58	Alternative	83	Attempt
9	Inverse	34	Manual	59	Abstract	84	Access
10	Partition	35	Manifest	60	Value	85	Term
11	Disproportionately	36	Link	61	Standard	86	Support
12	Requisite	37	Institute	62	Result	87	Subject
13	Imperative	38	Input	63	Research	88	Shaping
14	Diffuse	39	Index	64	Regard	89	Quarterly
15	Differential	40	Import	65	Reference	90	Present
16	Composite	41	Ideal	66	Range	91	Group
17	Patent	42	Function	67	Progressive	92	General
18	Endeavor	43	Fragment	68	Process	93	Form
19	Critique	44	Focus	69	Model	94	Following
20	Contrary	45	Estimate	70	Measure	95	Experience
21	Complement	46	Encounter	71	Material	96	Dependent
22	Absent	47	Coordinate	72	Male	97	Control
23	Universal	48	Convert	73	Lack	98	Concern
24	Survey	49	Contrast	74	Influence	99	Collective
25	Structure	50	Conflict	75	Individual	100	Above

*Note.* Word 1-3 appear on the AVL in three different word classes, each. The rest appear in two different word classes.

### Appendix F

Words Appearing on the AVL (Gardner & Davies, 2014) More than Once which are also Produced in the Corpus in the Present Study, along with their Frequency

1	<u>Value</u> 197	14	Standard 6	27	Link 3	40	Dependent 2
2	Experience 23	15	<u>Conflict</u> 6	28	Form 3	41	Critique 2
3	<u>Author</u> 13	16	<u>Regard</u> 5	29	Design 3	42	Alternative 2
4	Individual 12	17	Control 5	30	Decision-Making 3	43	Radical 1
5	Result 11	18	<u>Contrast</u> 5	31	<u>Access</u> 3	44	Prior 1
6	Lack 11	19	Range 4	32	<u>Subject</u> 2	45	Measure 1
7	Support 10	20	<u>Multiple</u> 4	33	Potential 2	46	Manual 1
8	Present 9	21	Influence 4	34	Parallel 2	47	Ideal 1
9	<u>Term</u> 8	22	Increase 4	35	<u>Overall</u> 2	48	Female 1
10	<u>Effect</u> 8	23	<u>Group</u> 4	36	<u>Material</u> 2	49	Contrary 1
11	<u>Concern</u> 8	24	<u>Appropriate</u> 4	37	<u>Function</u> 2	50	Attempt 1
12	<u>Following</u> 7	25	<u>Reference</u> 3	38	Estimate 2		
13	Focus 7	26	<u>Process</u> 3	39	Encounter 2		

*Note.* Underlined words appear in one POS within the top 500 lemmas on the AVL, and in one POS outside the top 500.

## Appendix G

### Complete Overview of AVL Types Produced in the Corpus and their Frequency

1	Value 197	128	Changing 5	255	Material 2	382	Mainstream 1
2	Example 73	129	Analysis 5	256	Limit 2	383	Leisure 1
3	Article 61	130	Address 5	257	Labor 2	384	Lastly 1
4	Technology 57	131	Viable 4	258	Justify 2	385	Judgment 1
5	Important 55	132	Use 4	259	Introduction 2	386	Judgement 1
6	Change 50	133	Transformation 4	260	Interact 2	387	Isolate 1
7	However 39	134	Tradition 4	261	Insight 2	388	Intolerance 1
8	Belief 38	135	Stress 4	262	Incorporate 2	389	International 1
9	Society 32	136	Role 4	263	Improvement 2	390	Interestingly 1
10	Describe 28	137	Represent 4	264	Household 2	391	Intention 1
11	Social 27	138	Relationship 4	265	Growth 2	392	Integrity 1
12	Therefore 26	139	Range 4	266	Function 2	393	Integrate 1
13	Human 25	140	Promote 4	267	Frequently 2	394	Intangible 1
14	Language 24	141	Profession 4	268	Figure 2	395	Initial 1
15	Experience 23	142	Obligation 4	269	Explicit 2	396	Inhabitant 1
16	Culture 22	143	Nonetheless 4	270	Evolution 2	397	Informative 1
17	Positive 21	144	Nevertheless 4	271	Ethnicity 2	398	Inform 1
18	Express 21	145	Multiple 4	272	Ethical 2	399	Infinite 1
19	Source 20	146	Manner 4	273	Ethic 2	400	Indicate 1
20	Problematic 20	147	Low 4	274	Estimate 2	401	Independent 1
21	Task 19	148	Involve 4	275	Encounter 2	402	Increasingly 1
22	Negative 19	149	Influence 4	276	Emotional 2	403	Inconsistent 1
23	Modern 19	150	Increase 4	277	Effective 2	404	Incompatible 1
24	History 19	151	Importance 4	278	Disregard 2	405	Inclusive 1
25	Instance 15	152	Image 4	279	Discussion 2	406	Imagery 1
26	Information 15	153	Highlight 4	280	Disagreement 2	407	Ideal 1
27	Include 15	154	Group 4	281	Determine 2	408	Historical 1
28	Current 15	155	Goal 4	282	Dependent 2	409	Helpful 1
29	View 13	156	Future 4	283	Demonstrate 2	410	Harvesting 1
30	Develop 13	157	Discrimination 4	284	Decline 2	411	Guiding 1
31	Content 13	158	Dilemma 4	285	Cycle 2	412	Guideline 1
32	Author 13	159	Description 4	286	Critique 2	413	Globalization 1
33	Argue 13	160	Convey 4	287	Consensus 2	414	Fundamental 1
34	Affect 13	161	Conclusion 4	288	Connection 2	415	Fulfil 1
35	Type 12	162	Appropriate 4	289	Compatible 2	416	Formulate 1
36	Tool 12	163	Achieve 4	290	Basis 2	417	Format 1
37	Tend 12	164	Word 3	291	Assume 2	418	Focused 1
38	Solution 12	165	Variety 3	292	Associate 2	419	Fluctuate 1
39	Similar 12	166	Valuable 3	293	Assimilate 2	420	Flow 1
40	Individual 12	167	Unrelated 3	294	Apply 2	421	Female 1
41	Extent 12	168	University 3	295	Alternative 2	422	Failure 1
42	Encourage 12	169	Trend 3	296	Adulthood 2	423	Extract 1
43	Basic 12	170	Trait 3	297	Accrue 2	424	Extend 1
44	Result 11	171	System 3	298	Accessible 2	425	Exploit 1
45	Lack 11	172	Subjective 3	299	Academic 2	426	Exemplify 1
46	Common 11	173	Strongly 3	300	Willingness 1	427	Exclusively 1
47	Both 11	174	Stereotype 3	301	Whole 1	428	Exclusion 1
48	Base 11	175	Shape 3	302	Well-Developed 1	429	Exclude 1
49	Support 10	176	Restrict 3	303	Well-Being 1	430	Excessive 1
50	Deem 10	177	Relate 3	304	Waste 1	431	Exception 1
51	Context 10	178	Reflection 3	305	Voluntary 1	432	Evolving 1
52	Century 10	179	Reference 3	306	Vision 1	433	Evident 1
53	Attitude 10	180	Racial 3	307	Valid 1	434	Evaluate 1
54	Western 9	181	Provide 3	308	User 1	435	Etc 1
55	Refer 9	182	Process 3	309	Unimportant 1	436	Equivalent 1
56	Present 9	183	Practice 3	310	Typically 1	437	Entity 1
57	Knowledge 9	184	Potentially 3	311	Typical 1	438	Ensure 1

58	Generally 9	185	Particularly 3	312	Two-Way 1	439	Enrich 1
59	Various 8	186	Participate 3	313	Traditionally 1	440	Enquiry 1
60	Understanding 8	187	Paraphrase 3	314	Traditional 1	441	Encapsulate 1
61	Term 8	188	Occupation 3	315	Time-Consuming 1	442	Emphasize 1
62	State 8	189	Norm 3	316	Territory 1	443	Embody 1
63	Response 8	190	Natural 3	317	Tendency 1	444	Element 1
64	Reflect 8	191	Movement 3	318	Technological 1	445	Elaborate 1
65	Publish 8	192	Modify 3	319	Tangible 1	446	Efficiently 1
66	Previous 8	193	Mental 3	320	Table 1	447	Effectively 1
67	Population 8	194	Mean 3	321	Summary 1	448	Economics 1
68	Perspective 8	195	Literature 3	322	Successful 1	449	Economically 1
69	Need 8	196	Link 3	323	Striking 1	450	Dominate 1
70	Known 8	197	Interest 3	324	Stem 1	451	Divergence 1
71	Explore 8	198	Initially 3	325	Specific 1	452	Disturbance 1
72	Evolve 8	199	Indeed 3	326	Socially 1	453	Discriminating 1
73	Everyday 8	200	Improve 3	327	Simultaneously 1	454	Discriminate 1
74	Environment 8	201	Illustrate 3	328	Significance 1	455	Disadvantage 1
75	Effect 8	202	Identify 3	329	Short 1	456	Direct 1
76	Device 8	203	Highly 3	330	Shared 1	457	Difficulty 1
77	Development 8	204	Guide 3	331	Sensitivity 1	458	Differentiate 1
78	Cultural 8	205	Form 3	332	Segregation 1	459	Differ 1
79	Concern 8	206	Equality 3	333	Scale 1	460	Detrimental 1
80	Adapt 8	207	Engage 3	334	Rigidity 1	461	Detect 1
81	Within 7	208	Dynamic 3	335	Reveal 1	462	Detailed 1
82	Shift 7	209	Directly 3	336	Restricted 1	463	Desirable 1
83	Occur 7	210	Design 3	337	Researcher 1	464	Descendant 1
84	Impact 7	211	Decision-Making 3	338	Require 1	465	Degree 1
85	Furthermore 7	212	Data 3	339	Reproduce 1	466	Definition 1
86	Following 7	213	Critical 3	340	Repetitive 1	467	Database 1
87	Focus 7	214	Creative 3	341	Reliable 1	468	Cultivate 1
88	Difference 7	215	Contact 3	342	Regardless 1	469	Crucial 1
89	Consequence 7	216	Consideration 3	343	Reduce 1	470	Counterproductive 1
90	Useful 6	217	Conclude 3	344	Recourse 1	471	Core 1
91	Theory 6	218	Comparison 3	345	Recognition 1	472	Cooperation 1
92	Theme 6	219	Communication 3	346	Realization 1	473	Contribution 1
93	Suggest 6	220	Cite 3	347	Re-Evaluate 1	474	Contrary 1
94	Standard 6	221	Challenge 3	348	Rationalize 1	475	Conspicuous 1
95	Rely 6	222	Access 3	349	Rapidly 1	476	Consequently 1
96	Relevant 6	223	Acceptable 3	350	Radical 1	477	Conform 1
97	Principle 6	224	Virtue 2	351	Protection 1	478	Conclusively 1
98	Meaning 6	225	Viewpoint 2	352	Prominent 1	479	Compulsory 1
99	Interpret 6	226	Unrealistic 2	353	Profound 1	480	Comprehend 1
100	Harmful 6	227	Transform 2	354	Profile 1	481	Commitment 1
101	Firstly 6	228	Tolerance 2	355	Professionalism 1	482	Commentary 1
102	Factor 6	229	Subject 2	356	Professional 1	483	Combination 1
103	Facility 6	230	Study 2	357	Production 1	484	Center 1
104	Conflict 6	231	Strive 2	358	Product 1	485	Category 1
105	Aspect 6	232	Societal 2	359	Prioritize 1	486	Categorize 1
106	Verbatim 5	233	Socialize 2	360	Prior 1	487	Broad 1
107	Regard 5	234	Similarity 2	361	Previously 1	488	Beneficial 1
108	Purpose 5	235	Separate 2	362	Preservation 1	489	Available 1
109	Prejudice 5	236	Seek 2	363	Poverty 1	490	Attempt 1
110	Period 5	237	Scenario 2	364	Positively 1	491	Association 1
111	Negatively 5	238	Report 2	365	Perpetuate 1	492	Altered 1
112	Likely 5	239	Relation 2	366	Perceive 1	493	Alteration 1
113	Level 5	240	Quality 2	367	Pattern 1	494	Align 1
114	Global 5	241	Potential 2	368	Passage 1	495	Aid 1
115	Exist 5	242	Philosophy 2	369	Originate 1	496	Advanced 1
116	Educated 5	243	Phenomenon 2	370	Occasion 1	497	Adult 1
117	Diversity 5	244	Personalized 2	371	Obtain 1	498	Adopt 1



<b>118</b>	Discuss 5	<b>245</b>	Perform 2	<b>372</b>	Noticeable 1	<b>499</b>	Admittedly 1
<b>119</b>	Developing 5	<b>246</b>	Parallel 2	<b>373</b>	Note 1	<b>500</b>	Additional 1
<b>120</b>	Depend 5	<b>247</b>	Paragraph 2	<b>374</b>	Multitude 1	<b>501</b>	Accordingly 1
<b>121</b>	Control 5	<b>248</b>	Overall 2	<b>375</b>	Multifaceted 1	<b>502</b>	Acceptance 1
<b>122</b>	Contribute 5	<b>249</b>	Organize 2	<b>376</b>	Motivate 1	<b>503</b>	Absolute 1
<b>123</b>	Contrast 5	<b>250</b>	Numerous 2	<b>377</b>	Moreover 1	<b>504</b>	Aberrant 1
<b>124</b>	Contain 5	<b>251</b>	Neutral 2	<b>378</b>	Method 1	<b>505</b>	Abbreviation 1
<b>125</b>	Concept 5	<b>252</b>	Neglect 2	<b>379</b>	Measure 1		
<b>126</b>	Complex 5	<b>253</b>	Necessary 2	<b>380</b>	Manual 1		
<b>127</b>	Compare 5	<b>254</b>	Merely 2	<b>381</b>	Maintain 1		

---

## Appendix H

### AVL Token Coverage of the Corpus in Percentages

<b>Learner ID</b>	<b>Task 1</b>	<b>Task 2</b>	<b>Task 3</b>	<b>Total</b>
<i>AA</i>	6,25 %	1,51 %	7,86 %	5,80 %
<i>AB</i>	10,45 %	5,02 %	3,78 %	5,42 %
<i>AC</i>	11,11 %	6,23 %	10,42 %	9,53 %
<i>AD</i>	7,61 %	2,78 %	4,49 %	4,65 %
<i>AE</i>	10,90 %	8,43 %	18,50 %	15,13 %
<i>AF</i>	7,25 %	4,44 %	7,38 %	6,74 %
<i>AG</i>	4,52 %	5,98 %	0,43 %	2,31 %
<i>AH</i>	6,36 %	9,06 %	17,15 %	13,04 %
<i>AI</i>	7,96 %	6,39 %	9,38 %	8,42 %
<i>AJ</i>	11,42 %	5,02 %	2,96 %	5,37 %
<i>AK</i>	13,47 %	6,43 %	1,90 %	4,75 %
<i>AL</i>	4,98 %	1,92 %	10,95 %	8,23 %
<i>AM</i>	10,00 %	5,49 %	9,94 %	8,90 %
<i>AN</i>	15,05 %	9,02 %	11,49 %	11,61 %
<i>AO</i>	4,56 %	5,77 %	9,54 %	7,85 %
<i>AP</i>	8,81 %	5,71 %	13,05 %	10,73 %
<i>AQ</i>	7,98 %	2,63 %	21,95 %	15,92 %
<i>AR</i>	4,62 %	3,43 %	6,23 %	5,31 %
<i>AS</i>	6,76 %	6,11 %	6,28 %	6,33 %
<i>AT</i>	2,35 %	4,36 %	4,29 %	3,99 %
<i>AU</i>	9,55 %	2,75 %	10,37 %	8,58 %
<i>AV</i>	11,11 %	6,50 %	11,65 %	10,00 %
<i>BA</i>	9,04 %	7,41 %	11,91 %	10,22 %
<i>BB</i>	6,47 %	8,57 %	5,33 %	5,96 %
<i>BC</i>	13,00 %	8,29 %	8,75 %	9,33 %
<i>BD</i>	6,84 %	12,44 %	8,53 %	8,91 %
<i>BE</i>	3,00 %	6,63 %	6,44 %	5,87 %
<i>BF</i>	4,62 %	3,03 %	4,21 %	4,04 %
<b>Total</b>	7,98 %	5,72 %	8,67 %	7,91 %

### Appendix I

#### AVL Token Coverage of the Corpus in Raw Numbers

<b>Learner ID</b>	<b>Task 1</b>	<b>Task 2</b>	<b>Task 3</b>	<b>Total</b>
<i>AA</i>	13	4	40	57
<i>AB</i>	21	13	21	55
<i>AC</i>	21	16	64	101
<i>AD</i>	15	7	31	53
<i>AE</i>	17	21	128	166
<i>AF</i>	15	11	54	80
<i>AG</i>	9	14	3	26
<i>AH</i>	11	25	94	130
<i>AI</i>	16	17	61	94
<i>AJ</i>	29	11	18	58
<i>AK</i>	26	16	14	56
<i>AL</i>	13	3	74	90
<i>AM</i>	17	14	65	96
<i>AN</i>	31	22	71	124
<i>AO</i>	11	15	75	101
<i>AP</i>	17	12	80	109
<i>AQ</i>	13	7	176	196
<i>AR</i>	9	8	39	56
<i>AS</i>	15	19	37	71
<i>AT</i>	5	13	35	53
<i>AU</i>	19	7	75	101
<i>AV</i>	18	16	48	82
<i>BA</i>	17	14	53	84
<i>BB</i>	13	18	55	86
<i>BC</i>	26	16	81	123
<i>BD</i>	13	25	71	109
<i>BE</i>	6	12	50	68
<i>BF</i>	9	6	21	36
<b>Total</b>	445	382	1,634	2,461

### Appendix J

Distribution of AVL Types in the Corpus in Raw Numbers

<b>Learner ID</b>	<b>Task 1</b>	<b>Task 2</b>	<b>Task 3</b>	<b>Total</b>
<i>AA</i>	13	4	25	36
<i>AB</i>	18	13	18	44
<i>AC</i>	15	14	36	52
<i>AD</i>	12	7	22	35
<i>AE</i>	15	21	69	96
<i>AF</i>	12	11	30	49
<i>AG</i>	6	13	3	21
<i>AH</i>	11	20	54	79
<i>AI</i>	12	16	35	59
<i>AJ</i>	17	10	16	39
<i>AK</i>	15	13	12	36
<i>AL</i>	9	3	46	56
<i>AM</i>	16	14	45	66
<i>AN</i>	27	21	52	93
<i>AO</i>	9	15	52	70
<i>AP</i>	13	12	34	54
<i>AQ</i>	12	7	90	106
<i>AR</i>	5	8	18	28
<i>AS</i>	12	15	25	47
<i>AT</i>	5	9	22	33
<i>AU</i>	14	6	45	63
<i>AV</i>	16	14	32	56
<i>BA</i>	12	11	24	44
<i>BB</i>	9	14	36	57
<i>BC</i>	23	13	52	83
<i>BD</i>	9	19	40	59
<i>BE</i>	6	9	36	41
<i>BF</i>	6	5	16	26
<b>Total</b>	154	177	402	505

## Appendix K

Types Covered by the First 500 Lemmas on the AVL Including their Frequency of

Occurrence in the Corpus

1	Value 197	73	Concern 8	145	Provide 3	217	Associate 2
2	Example 73	74	Occur 7	146	Process 3	218	Apply 2
3	Article 61	75	Impact 7	147	Practice 3	219	Alternative 2
4	Technology 57	76	Furthermore 7	148	Particularly 3	220	Waste 1
5	Important 55	77	Following 7	149	Participate 3	221	Vision 1
6	Change 50	78	Focus 7	150	Natural 3	222	User 1
7	However 39	79	Difference 7	151	Movement 3	223	Typically 1
8	Belief 38	80	Consequence 7	152	Mental 3	224	Typical 1
9	Society 32	81	Useful 6	153	Mean 3	225	Traditional 1
10	Describe 28	82	Theory 6	154	Literature 3	226	Territory 1
11	Social 27	83	Theme 6	155	Link 3	227	Table 1
12	Therefore 26	84	Suggest 6	156	Interest 3	228	Successful 1
13	Human 25	85	Standard 6	157	Indeed 3	229	Specific 1
14	Language 24	86	Rely 6	158	Improve 3	230	Significance 1
15	Experience 23	87	Relevant 6	159	Illustrate 3	231	Scale 1
16	Culture 22	88	Principle 6	160	Identify 3	232	Reveal 1
17	Positive 21	89	Meaning 6	161	Highly 3	233	Researcher 1
18	Express 21	90	Interpret 6	162	Form 3	234	Require 1
19	Source 20	91	Factor 6	163	Engage 3	235	Reduce 1
20	Task 19	92	Facility 6	164	Directly 3	236	Recognition 1
21	Negative 19	93	Conflict 6	165	Design 3	237	Protection 1
22	Modern 19	94	Aspect 6	166	Data 3	238	Professional 1
23	History 19	95	Regard 5	167	Critical 3	239	Production 1
24	Information 15	96	Purpose 5	168	Creative 3	240	Product 1
25	Include 15	97	Period 5	169	Contact 3	241	Previously 1
26	Current 15	98	Likely 5	170	Consideration 3	242	Perceive 1
27	View 13	99	Level 5	171	Conclude 3	243	Pattern 1
28	Develop 13	100	Global 5	172	Comparison 3	244	Passage 1
29	Content 13	101	Exist 5	173	Communication 3	245	Obtain 1
30	Author 13	102	Diversity 5	174	Cite 3	246	Note 1
31	Argue 13	103	Discuss 5	175	Challenge 3	247	Moreover 1
32	Affect 13	104	Developing 5	176	Access 3	248	Method 1
33	Type 12	105	Depend 5	177	Transform 2	249	Measure 1
34	Tool 12	106	Control 5	178	Subject 2	250	Maintain 1
35	Tend 12	107	Contribute 5	179	Study 2	251	Judgment 1
36	Solution 12	108	Contrast 5	180	Separate 2	252	International 1
37	Similar 12	109	Contain 5	181	Seek 2	253	Initial 1
38	Individual 12	110	Concept 5	182	Report 2	254	Inform 1
39	Extent 12	111	Complex 5	183	Relation 2	255	Indicate 1
40	Encourage 12	112	Compare 5	184	Quality 2	256	Independent 1
41	Basic 12	113	Analysis 5	185	Potential 2	257	Increasingly 1
42	Result 11	114	Address 5	186	Philosophy 2	258	Historical 1
43	Lack 11	115	Use 4	187	Phenomenon 2	259	Fundamental 1

<b>44</b> Common 11	<b>116</b> Tradition 4	<b>188</b> Perform 2	<b>260</b> Flow 1
<b>45</b> Both 11	<b>117</b> Role 4	<b>189</b> Overall 2	<b>261</b> Female 1
<b>46</b> Base 11	<b>118</b> Represent 4	<b>190</b> Organize 2	<b>262</b> Failure 1
<b>47</b> Support 10	<b>119</b> Relationship 4	<b>191</b> Numerous 2	<b>263</b> Extend 1
<b>48</b> Context 10	<b>120</b> Range 4	<b>192</b> Necessary 2	<b>264</b> Evaluate 1
<b>49</b> Century 10	<b>121</b> Promote 4	<b>193</b> Merely 2	<b>265</b> Ensure 1
<b>50</b> Attitude 10	<b>122</b> Nevertheless 4	<b>194</b> Material 2	<b>266</b> Emphasize 1
<b>51</b> Western 9	<b>123</b> Multiple 4	<b>195</b> Limit 2	<b>267</b> Element 1
<b>52</b> Refer 9	<b>124</b> Manner 4	<b>196</b> Labor 2	<b>268</b> Effectively 1
<b>53</b> Present 9	<b>125</b> Low 4	<b>197</b> Introduction 2	<b>269</b> Direct 1
<b>54</b> Knowledge 9	<b>126</b> Involve 4	<b>198</b> Insight 2	<b>270</b> Difficulty 1
<b>55</b> Generally 9	<b>127</b> Influence 4	<b>199</b> Incorporate 2	<b>271</b> Differ 1
<b>56</b> Various 8	<b>128</b> Increase 4	<b>200</b> Improvement 2	<b>272</b> Degree 1
<b>57</b> Understanding 8	<b>129</b> Importance 4	<b>201</b> Household 2	<b>273</b> Definition 1
<b>58</b> Term 8	<b>130</b> Image 4	<b>202</b> Growth 2	<b>274</b> Crucial 1
<b>59</b> State 8	<b>131</b> Group 4	<b>203</b> Function 2	<b>275</b> Core 1
<b>60</b> Response 8	<b>132</b> Goal 4	<b>204</b> Frequently 2	<b>276</b> Cooperation 1
<b>61</b> Reflect 8	<b>133</b> Future 4	<b>205</b> Figure 2	<b>277</b> Contribution 1
<b>62</b> Publish 8	<b>134</b> Description 4	<b>206</b> Evolution 2	<b>278</b> Commitment 1
<b>63</b> Previous 8	<b>135</b> Conclusion 4	<b>207</b> Estimate 2	<b>279</b> Combination 1
<b>64</b> Population 8	<b>136</b> Appropriate 4	<b>208</b> Emotional 2	<b>280</b> Category 1
<b>65</b> Perspective 8	<b>137</b> Achieve 4	<b>209</b> Effective 2	<b>281</b> Broad 1
<b>66</b> Need 8	<b>138</b> Variety 3	<b>210</b> Discussion 2	<b>282</b> Available 1
<b>67</b> Explore 8	<b>139</b> University 3	<b>211</b> Determine 2	<b>283</b> Attempt 1
<b>68</b> Environment 8	<b>140</b> Trend 3	<b>212</b> Demonstrate 2	<b>284</b> Association 1
<b>69</b> Effect 8	<b>141</b> System 3	<b>213</b> Cycle 2	<b>285</b> Adopt 1
<b>70</b> Device 8	<b>142</b> Strongly 3	<b>214</b> Connection 2	<b>286</b> Additional 1
<b>71</b> Development 8	<b>143</b> Relate 3	<b>215</b> Basis 2	
<b>72</b> Cultural 8	<b>144</b> Reference 3	<b>216</b> Assume 2	

---

## Appendix L

AVL Types Produced in the Corpus Found in the Assignment Briefs.

1	Use	51	Current	101	Manner
2	Task	52	Conflict	102	Level
3	Need	53	Compare	103	Knowledge
4	Value	54	Common	104	Involve
5	Example	55	Basic	105	Information
6	Discuss	56	Author	106	Acceptable
7	Change	57	Affect	107	Individual
8	Source	58	Adapt	108	Indeed
9	Limit	59	Academic	109	Increasingly
10	Language	60	Whole	110	Improve
11	Include	61	Access	111	Important
12	Future	62	Viable	112	Impact
13	Type	63	Verbatim	113	Identify
14	Table	64	Useful	114	However
15	Response	65	University	115	Appropriate
16	Article	66	Term	116	Harmful
17	Technology	67	Suggest	117	Guide
18	Modern	68	Strive	118	Goal
19	Mean	69	Society	119	Focus
20	Material	70	Similar	120	Explore
21	Interest	71	Short	121	Exist
22	Human	72	Shape	122	Evolution
23	Tool	73	Result	123	Evident
24	Support	74	Researcher	124	Everyday
25	Social	75	Require	125	Etc
26	Experience	76	Report	126	Encounter
27	Directly	77	Rely	127	Dynamic
28	Describe	78	Relationship	128	Discussion
29	Cite	79	Reflection	129	Difficulty
30	Western	80	Refer	130	Differ
31	Achieve	81	Racial	131	Detect
32	Study	82	Analysis	132	Design
33	Reference	83	Professional	133	Description
34	Provide	84	Profession	134	Depend
35	Problematic	85	Previously	135	Decline
36	Positive	86	Previous	136	Cultural
37	Literature	87	Present	137	Critique
38	Lack	88	Practice	138	Convey
39	Interpret	89	Potentially	139	Aspect
40	Influence	90	Perspective	140	Context
41	Image	91	Personalized	141	Content

<b>42</b>	Group	<b>92</b>	Period	<b>142</b>	Contain
<b>43</b>	Global	<b>93</b>	Particularly	<b>143</b>	Century
<b>44</b>	Generally	<b>94</b>	Paraphrase	<b>144</b>	Both
<b>45</b>	Extent	<b>95</b>	Parallel	<b>145</b>	Belief
<b>46</b>	Express	<b>96</b>	Organize	<b>146</b>	Attitude
<b>47</b>	Engage	<b>97</b>	Obligation	<b>147</b>	Adulthood
<b>48</b>	Device	<b>98</b>	Note	<b>148</b>	Following
<b>49</b>	Development	<b>99</b>	Negative	<b>149</b>	Decision-making
<b>50</b>	Develop	<b>100</b>	Meaning		

---



### Appendix M

AVL Types Produced in the Corpus Found in the Assignment Briefs and Preparation Material  
Combined.

<b>1</b>	Use	<b>82</b>	Analysis	<b>163</b>	Adulthood
<b>2</b>	Language	<b>83</b>	Advanced	<b>164</b>	Purpose
<b>3</b>	Change	<b>84</b>	Whole	<b>165</b>	Promote
<b>4</b>	Need	<b>85</b>	Western	<b>166</b>	Professional
<b>5</b>	Value	<b>86</b>	Typically	<b>167</b>	Profession
<b>6</b>	Include	<b>87</b>	Role	<b>168</b>	Product
<b>7</b>	Information	<b>88</b>	Relevant	<b>169</b>	Process
<b>8</b>	Future	<b>89</b>	Relate	<b>170</b>	Principle
<b>9</b>	Task	<b>90</b>	Regardless	<b>171</b>	Previously
<b>10</b>	Human	<b>91</b>	Reflect	<b>172</b>	Potentially
<b>11</b>	Century	<b>92</b>	Racial	<b>173</b>	Personalized
<b>12</b>	Source	<b>93</b>	Publish	<b>174</b>	Period
<b>13</b>	Social	<b>94</b>	Problematic	<b>175</b>	Perform
<b>14</b>	Modern	<b>95</b>	Present	<b>176</b>	Particularly
<b>15</b>	Interest	<b>96</b>	Practice	<b>177</b>	Paraphrase
<b>16</b>	Experience	<b>97</b>	Passage	<b>178</b>	Parallel
<b>17</b>	Discuss	<b>98</b>	Organize	<b>179</b>	Paragraph
<b>18</b>	Author	<b>99</b>	Occur	<b>180</b>	Obligation
<b>19</b>	View	<b>100</b>	Note	<b>181</b>	Necessary
<b>20</b>	Theory	<b>101</b>	Lack	<b>182</b>	Natural
<b>21</b>	Technology	<b>102</b>	Interpret	<b>183</b>	Motivate
<b>22</b>	Limit	<b>103</b>	International	<b>184</b>	Moreover
<b>23</b>	Knowledge	<b>104</b>	Influence	<b>185</b>	Modify
<b>24</b>	Group	<b>105</b>	Individual	<b>186</b>	Meaning
<b>25</b>	Global	<b>106</b>	Increasingly	<b>187</b>	Manner
<b>26</b>	Example	<b>107</b>	Improve	<b>188</b>	Maintain
<b>27</b>	Both	<b>108</b>	History	<b>189</b>	Mainstream
<b>28</b>	Type	<b>109</b>	Guide	<b>190</b>	Low
<b>29</b>	Table	<b>110</b>	Female	<b>191</b>	Labor
<b>30</b>	Provide	<b>111</b>	Extent	<b>192</b>	Apply
<b>31</b>	Material	<b>112</b>	Express	<b>193</b>	Indeed
<b>32</b>	Image	<b>113</b>	Estimate	<b>194</b>	Increase
<b>33</b>	Article	<b>114</b>	Ensure	<b>195</b>	Improvement
<b>34</b>	Term	<b>115</b>	Encounter	<b>196</b>	Importance
<b>35</b>	Support	<b>116</b>	Emphasize	<b>197</b>	Impact
<b>36</b>	Response	<b>117</b>	Effect	<b>198</b>	Identify
<b>37</b>	Reference	<b>118</b>	Diversity	<b>199</b>	However
<b>38</b>	Mean	<b>119</b>	Difficulty	<b>200</b>	Harmful
<b>39</b>	Likely	<b>120</b>	Description	<b>201</b>	Guideline
<b>40</b>	Important	<b>121</b>	Depend	<b>202</b>	Goal

<b>41</b>	Engage	<b>122</b>	Cultural	<b>203</b>	Furthermore
<b>42</b>	Describe	<b>123</b>	Control	<b>204</b>	Fundamental
<b>43</b>	Tool	<b>124</b>	Content	<b>205</b>	Format
<b>44</b>	Result	<b>125</b>	Conflict	<b>206</b>	Form
<b>45</b>	Report	<b>126</b>	Conclude	<b>207</b>	Figure
<b>46</b>	Previous	<b>127</b>	Common	<b>208</b>	Failure
<b>47</b>	Literature	<b>128</b>	Challenge	<b>209</b>	Explore
<b>48</b>	Involve	<b>129</b>	Available	<b>210</b>	Evolution
<b>49</b>	Exist	<b>130</b>	Access	<b>211</b>	Evident
<b>50</b>	Directly	<b>131</b>	Adapt	<b>212</b>	Everyday
<b>51</b>	Device	<b>132</b>	Accessible	<b>213</b>	Ethnicity
<b>52</b>	Development	<b>133</b>	Academic	<b>214</b>	Ethic
<b>53</b>	Develop	<b>134</b>	Within	<b>215</b>	Etc
<b>54</b>	Design	<b>135</b>	Viewpoint	<b>216</b>	Emotional
<b>55</b>	Deem	<b>136</b>	Viable	<b>217</b>	Dynamic
<b>56</b>	Cite	<b>137</b>	Verbatim	<b>218</b>	Discussion
<b>57</b>	Basic	<b>138</b>	Valuable	<b>219</b>	Differ
<b>58</b>	Affect	<b>139</b>	University	<b>220</b>	Determine
<b>59</b>	Waste	<b>140</b>	Traditionally	<b>221</b>	Detect
<b>60</b>	Useful	<b>141</b>	Therefore	<b>222</b>	Degree
<b>61</b>	Technological	<b>142</b>	Tendency	<b>223</b>	Definition
<b>62</b>	Study	<b>143</b>	Suggest	<b>224</b>	Decline
<b>63</b>	Society	<b>144</b>	Strive	<b>225</b>	Culture
<b>64</b>	Require	<b>145</b>	Stereotype	<b>226</b>	Critique
<b>65</b>	Refer	<b>146</b>	Acceptable	<b>227</b>	Convey
<b>66</b>	Positive	<b>147</b>	Similar	<b>228</b>	Contrast
<b>67</b>	Perspective	<b>148</b>	Significance	<b>229</b>	Context
<b>68</b>	Neutral	<b>149</b>	Short	<b>230</b>	Contain
<b>69</b>	Negative	<b>150</b>	Shape	<b>231</b>	Consequently
<b>70</b>	Level	<b>151</b>	Sensitivity	<b>232</b>	Concern
<b>71</b>	Inclusive	<b>152</b>	Scenario	<b>233</b>	Complex
<b>72</b>	Highly	<b>153</b>	Achieve	<b>234</b>	Center
<b>73</b>	Highlight	<b>154</b>	Reveal	<b>235</b>	Broad
<b>74</b>	Generally	<b>155</b>	Researcher	<b>236</b>	Attitude
<b>75</b>	Focus	<b>156</b>	Represent	<b>237</b>	Attempt
<b>76</b>	Environment	<b>157</b>	Rely	<b>238</b>	Association
<b>77</b>	Current	<b>158</b>	Reliable	<b>239</b>	Aspect
<b>78</b>	Critical	<b>159</b>	Relationship	<b>240</b>	Appropriate
<b>79</b>	Compare	<b>160</b>	Reflection	<b>241</b>	Following
<b>80</b>	Communication	<b>161</b>	Rapidly	<b>242</b>	Data
<b>81</b>	Belief	<b>162</b>	Range	<b>243</b>	Decision-making

---

## Appendix N

### AVL Words (in Bold) Evaluated as Incorrect and the Assumed Intended Meaning

Sentence	Intended meaning
The reason they get so much <b>critique</b> is that they have the power to decide where the world cup will be held.	Criticism
The decision whether it is good to change the past is complex as it <b>effects</b> people from all sides.	Affects
This is proven by looking at the old <b>achieves</b> .	Archives
In <b>reference</b> , the text includes a passage dated back to 1925 expressing that the youth is «best described as grossly thoughtless, rude and utterly selfish”	(For) instance
Values are often passed down through family, <b>development</b> from own experiences as well as history.	Developed
On the other hand, some values that recently <b>accrued</b> have been sensed as aberrant to the already socially standing values.	Appeared or emerged
These laws may have <b>accrued</b> because of more nationalistic values that generations before us are more familiar to	Appeared or emerged
The values of a person will have great impact on the persons decisions, and therefore <b>effect</b> one’s life.	Affect
<b>Including</b> this there were also a lot of injuries.	In addition to
That is why it is super important to be well educated about these kinds of things, so that we are able to <b>provide</b> and avoid them from happening again.	Prevent
Another theme that the younger generations often receive <b>critique</b> about, is their spending.	Criticism
We are on daily basis <b>effected</b> on technology which makes our life easier.	Dependent or affected by
During life we <b>obtain</b> a ton of values that change and disappears.	Have or acquire
It is when he becomes enamored with Lord Henry’s witty hedonistic ideas, that his new values <b>conclusively</b> transform him.	Consequently or ultimately
For instance, has several of the well-known children’s books written by author Roald Dahl been rewritten to be <b>viable</b> to today’s values and believes.	Acceptable or appropriate
However, I do not think it is either correct or respectful of you to express that untidiness at school result in students not having what it takes to become the professions you are <b>illustrating</b> .	Highlighting
Through the text we notice the <b>repetitive</b> comparison between the modern-day young and the old people when they were younger.	Repeated
When we are young our values often <b>represent</b> our parents’ values, but as we go through life, we may encounter new situations or people that challenge our existing values or beliefs.	Reflect

---

One must always think twice before <b>expressing</b> .	Expressing oneself or speaking
Lastly, one should also make sure one does not recycle <b>previous</b> submitted works, as this is seen as self-plagiarism.	Previously

---

### Appendix O

#### Descriptive Statistics on AVL Vocabulary from the Assignment Briefs and Preparation

##### Material in the Corpus Based on Grades

	Mid group	High group	Total
All tokens	16,321	14,788	31,109
All AVL tokens	1,199	1,262	2,461
AVL tokens from assignment briefs	782	774	1,556
Proportion of AVL tokens made up from AVL types in the assignment briefs	65.22%	61.33%	63.23%
AVL tokens from assignment briefs + preparation material	947	946	1,893
Proportion of AVL tokens made up from AVL types in the assignment briefs + preparation material	78.98%	74.96%	76.92%

### Appendix P

Descriptive Statistics on AVL Use Based on Grades and the Tokens Made up from the Top  
500 Lemmas on the AVL by Gardner and Davies (2014)

	Mid group	High group	Total
All tokens	16,321	14,788	31,109
All AVL tokens	1,199	1,262	2,461
Top 500 AVL tokens	1,005	1,020	2,025
Proportion of AVL tokens made up from top 500 AVL	83.82%	80.82%	82.28%
Coverage of all tokens provided by top 500 AVL	6.16%	6.90%	6.51%