Lexical Errors in Norwegian Intermediate and Advanced Learners of English

Karina Rose Mahan

A thesis presented to
The Department of Literature, Area Studies and European Languages

UNIVERSITY OF OSLO

SPRING 2013

Abstract

This thesis builds on previous taxonomies of lexical errors in order to find patterns of lexical errors in Norwegian intermediate and advanced English texts. This study uses a taxonomy created by Angela Hasselgren to determine patterns of lexical errors. Lexical errors are labeled by their routes (how the learner chooses the wrong word), effects (why the word is wrong) and influences (what causes the error). By locating and labeling lexical errors in two corpora, this study illustrates how lexical errors are distributed across texts written by Norwegian speakers of English. The distribution of errors uncovers patterns that in turn explain how and why some errors are repeated. By contrasting intermediate and advanced speakers, it is determined that the distributions of intermediate and advanced lexical errors are significantly different. The amount of errors is significantly smaller in the advanced texts. Comparative distributions show that intermediate and advanced learners have different ways of choosing words (routes), but the distribution of effects remains constant. Advanced learners appear to use direct L1-influence and intralingual influence more, while intermediate learners depend on indirect L1-influence.

Acknowledgements

This thesis would never have been written if it weren't for the help and support of many wonderful people. First I would like to thank Janne Skjelbred, who helped me compile the intermediate corpus. Without her help, there wouldn't have been an intermediate corpus to base my thesis on. I would also like to thank Lisbeth Brevik for her support, encouragement and much-appreciated suggestions.

A special thanks goes to my supervisor, Signe Oksefjell Ebeling. She gave invaluable advice on many subjects, helped me on very short notice, worked with extreme deadlines and thoroughly read and corrected this thesis. I would also like to thank my friends Eric and Aurelija for their technical support on word formatting and statistical analysis. Additionally I would like to thank my friend Myriam for thoroughly proofreading my thesis on an extremely short notice.

I found the process of writing a thesis often times demanding, but much less so in the company of fantastic neighbors, family and friends. Thanks to my neighbors who always had time for a cup and an ear to lend. I would especially like to thank my mother and Aurelija for their continual support and patience.

Table of Contents

Figures

List of Figures

Tables

List of tables

List of abbreviations

EFL – English as a Foreign Language

ICLE – International Corpus of Learner English

L1 – first language

L2 – second language

NICLE – the Norwegian part of the International Corpus of Learner English

PE – perceived equivalent

SLA – second language acquisition

1. INTRODUCTION

Language is traditionally divided into grammar and lexis. In second language acquisition/learning, they take the form of rule-learning (grammar) and vocabulary learning (lexis). Both are important to be able to communicate, yet second language research focuses much more on grammar. There is extensive research done on how learners develop grammatical patterns, what mistakes are typical of learners and how rules can be applied to prevent grammatical errors. Perhaps as a result of this, second language teaching puts much more emphasis on grammar than vocabulary in the second language classroom. This is done by explicitly teaching grammatical rules, and assuming that students will acquire vocabulary through osmosis. Most foreign language teachers have therefore been taught how to cope with grammatical errors (*They **is*** is a concord mistake), but not how to cope with lexical errors (*We had a **jolly nice** time* is a wrong word choice – but what makes it wrong?). Underdeveloped systems for vocabulary have large consequences for the EFL (English as a Foreign Language) classroom – teachers are left telling their student which word is right (*we had a jolly good time*), without offering a satisfactory explanation for why the other is wrong. Being able to explain **why** something is wrong is the first step to preventing similar errors. If systems for vocabulary are not developed, these mistakes can only be treated as singular errors that simply must be memorized.

Lexis has largely been ignored in SLA (second language acquisition) studies until the last 20-30 years. Researchers have blamed the "unstable and unsystematic nature of the lexicon" to justify their focus on grammatical aspects (Llach 2011: 70). To many, the lexicon is ostensibly a collection of unrelated, lone words that cannot be systematized. However, quantitative research on word frequencies can now confirm that words rarely occur alone in discourse. Psycholinguistics is also pointing toward a highly organized mental lexicon, where words are organized by both semantic (meaning-based) and formal aspects (such as similar word forms) (Aitchison 2012: 10). Such research suggests that a) words can be grouped together based on how they occur in discourse and b) words can be grouped together based on how they are acquired.

In SLA studies, techniques for identifying, classifying and systematically interpreting errors (also known as error analysis) are constantly being developed to account for the many errors that second language learners make (Crystal 2008: 173). Analysis of lexical errors is the first step to creating taxonomies of errors – i.e., systems that can explain how/why words are wrong. Quantitative studies of these mistakes in learner language can also determine which

mistakes are more reoccurring and characteristic of non-native speakers. This thesis is such a study. By examining texts written by Norwegian students of intermediate and advanced English, this study hopes to establish a) what the most reoccurring lexical mistakes in Norwegian speakers of English are and b) whether these reoccurring lexical mistakes are the same in both intermediate and advanced students.

1.1 Why vocabulary?

Vocabulary is the largest area of linguistic knowledge (Caspi & Lowie 2010: 46). Adult native English speakers with an educated background know around 15, 000 – 20, 000 *word families* (words with the same base, but belonging to different word classes, e.g. *beauty* - *beautiful*) (Schmitt 2010: 29). An L2 (second language) speaker must acquire approximately 3, 000 word families to minimally follow a conversation, and around 9, 000 word families to read novels or newspapers (Ortega 2009: 88). Unsurprisingly, lack of vocabulary does not only impede knowledge of the language – Meara comments that measurements of vocabulary size have been shown to correlate positively with proficiency levels in reading and writing, and in general language proficiency (Meara 2009: 34).

Vocabulary knowledge is two-sided – it is not only about **how many words you know** (vocabulary breadth), but also **how well you know them** (vocabulary depth). Studies in vocabulary acquisition have shown that learning vocabulary is not an instantaneous process – words are not either known/unknown, and knowledge of words most likely operates on a continuum (see section 3.3.1). For instance, Meara discovered that changes in response to words (e.g. moving from clang associations to native-speaker-like responses) still take place as much as twelve weeks after initial presentation of taught words (Meara 2009: 27). He also believes that knowing words is a "complex and multi-faceted skill, perhaps best described in behavioral terms as the ability to react to a word in ways which are considered appropriate by the speech community" (Meara 2009: 19). Many lexical researchers have noted that word knowledge works on several levels, from explicit levels (such as form-meaning linkage) to implicit levels (such as when it is appropriate to use a word and around whom). Word knowledge can also be receptive (speaker is only able to recognize a word) and productive (speakers can also produce word). Productive knowledge typically lags behind receptive knowledge by a few hundred words, and once enough knowledge of a word is acquired, it changes from being passive to active knowledge (Meara 2009: 30).

Perhaps due to a) the large number of words that need to be acquired to communicate

and b) the many levels and aspects of knowing a word, it has been shown that L2 students think a lack of vocabulary is one of their major problems in their production of English (Simensen 1998: 220).

1.2 Why lexical errors?

The L2 vocabulary, much like L2 grammar, is often riddled with mysterious mistakes that a native speaker would not make. Finding systems for these errors may explain how and why these mistakes occur. Error analysis researchers hope that by investigating these special L2 word errors, we might gain insight into how L2 speakers acquire words and why they use them in non-native ways. This is called the study of lexical errors, and examines how L2 speakers misuse or misconceive vocabulary. The importance of lexical errors can be highlighted by two findings:

1. Researchers have suggested that lexical errors are as much as three times more frequent than grammatical errors in second language learners (Blaas 1982, cited in Ellis 2008: 50).

2. Research has shown lexical errors cause more problems of interpretation and a higher degree of irritation to the native speaker (Johansson 1978: 71).

The problem with locating lexical errors is that researchers cannot agree on what lexical knowledge consists of. If we cannot establish what lexical knowledge is, how can we find systems to judge if someone has misunderstood the knowledge? Erroneously, many believe that since there are "so many words" in a language, it is impossible to say why picking a certain one will be wrong. Studies of lexical knowledge and errors can tell us that although there are (potentially) an infinite number of words and word combinations a speaker has at her/his disposal, the type of information connected to a word is finite (for instance, spelling, pronunciation, meaning, collocation and frequency), and therefore incorrect use of this information is also finite (an error of spelling, pronunciation, meaning, et cetera).

1.3 The current study

By applying methods from SLA (specifically error analysis) and corpus linguistics, and drawing on tenets of semantics and phraseology, this thesis will attempt to uncover distributional patterns of lexical errors in intermediate and advanced texts produced by Norwegian speakers of English. A basis for this study was presented in chapter 1. Chapter 2

provides the research questions which this thesis will answer. Llach points out that one can only find lexical errors based on a) what "knowing a word" implies and b) what counts as a word (Llach 2011: 72). Therefore chapter 3 is dedicated to clarify what a word is and what "knowing a word" entails. Chapter 4 will explain lexical errors and their sources, and chapter 5 will describe the taxonomy of lexical errors in this study. Chapter 6 consists of methods and materials. Chapter 7 is the data analysis, and chapter 8 will answer the research questions based on said analysis. Chapter 9 will conclude the main findings, suggest applicable uses for the findings in the EFL classroom and provide suggestions for further research. The data which the thesis bases itself on (the number and types of errors) can be found in full detail in appendices 3 (intermediate learner errors) and 4 (advanced learner errors).

2. RESEARCH QUESTIONS

This research paper is inspired by Angela Hasselgren's MA thesis (hovedoppgave) “Right Words, Wrong Words and Different Words: an investigation into the lexical coping of Norwegian advanced learners of English” (Hasselgren 1993). Hasselgren investigates how Norwegian university students cope with their L2 (English) lexicon by a) analyzing how the group selects wrong words in translation tasks, b) determining why the selected words are wrong and c) possible sources of these errors. The wrong words are *lexically dissonant*[1] because they create *dissonance* between the word and its context, rather than violating any rules of the language (Martin 1984: 130). Hasselgren develops a taxonomy for labeling different aspects of lexical dissonance in hopes that the taxonomy could explain all such dissonances based on certain, limited criteria (e.g. that there are five general ways a word can be wrong). Quantitative research of dissonances based on the taxonomy could aid language teaching by identifying what aspects of vocabulary L2 speakers struggle with the most, and diagnose exactly why an L2 vocabulary – even a grammatically flawless advanced learner's vocabulary – may strike a native speaker as “foreign” (Hasselgren 1993: 2). The lexical dissonances in Hasselgren's study are labeled by their **route** (why they are wrong), **effect** (how they are wrong) and **divergence** (the source of error, i.e. inter- and/or intralingual influence) (see section 5.2 for more information on routes, effects and divergence).

Hasselgren investigates dissonances in translation texts[2] written by Norwegian speakers of advanced English. The focus of this study is to see how lexical dissonance is manifested in **intermediate** speakers as well as advanced. Such a comparison may provide insight into if there are any differences in routes, effects and divergence between the intermediate and the advanced stage in Norwegian learners of English. By analyzing free production texts written by intermediate and advanced students, this study builds on or complements Hasselgren's study of lexical dissonance in Norwegian speakers of English.

1 The words *dissonance* and *dissonances* are used interchangeably in this thesis, based on whether it is one mistake (*a dissonance*) or if it refers to lexical errors in general (*dissonance*). Martin (1984) prefers the term *dissonances* to explain individual word errors (.e.g *one dissonance*), whereas Hasselgren (1993) uses *dissonance* as an uncountable term for the phenomenon, and uses *wrong word* in the case of countable.

2 Tasks where the L2 speaker translates from their L1 (Norwegian) to their L2 (English)

The main research question for this study is:

1. How is lexical dissonance manifested in free production texts written by Norwegian speakers of English (intermediate and advanced)?

A taxonomy for labeling lexical dissonance is provided, including routes, effects and influence. The main research question can thus be answered by several subquestions:

a) *How is lexical dissonance distributed across routes and effects in Norwegian intermediate and advanced learners of English?*

b) *If some routes and effects are larger than others, what are possible reasons for this?*

Each dissonance is assigned both a route (how it is wrong) and effect (why it is wrong). One subquestion of interest is therefore:

c) *How do routes and effects combine? Is there a discernible pattern between how a wrong word is chosen and why it is dissonant?*

These questions do not fully cover the question of influence, or source of dissonance. One aspect that should be examined is how large a role the first language plays in second language error production. There are several views on L1-influence on the target language. Ellis claims that "a large number – and in some cases perhaps most – of the errors that learners produce are intralingual in origin rather than transfer" (Ellis 2008: 55). A subquestion that can confirm or refute this statement is:

d) *What is the proportion of L1-influenced dissonances (interlingual) vs non-influenced L1-dissonances (intralingual)?*

Hasselgren's study has the following route distribution ranking: synonyms > transliterations > associations > cognates = perceived equivalents. Effects have the following distribution ranking: semantic > collocational > stylistic/connotational > invalid > syntactic. Since this study uses her taxonomy, it would be interesting to see if the categories hold the same rank in my data. Therefore, the last subquestion for the first research question is:

e) *Will my data have the same distributional patterns for routes and effects as Hasselgren's?*[3]

3 This is determined by a token ranking (see sections 6.3.2 and 7.1.1 for more information)

Since this study will analyze lexical dissonance at both the intermediate and advanced levels of Norwegian speakers of English, it is necessary to take into account that different levels of English may produce different distributions of dissonance, e.g. that intermediate students have different ways of choosing the wrong word than advanced students. The routes, effects and influence distributions are of course useful to look at from a wider perspective, but it is also necessary to investigate if there are differences between the intermediate and advanced learners. To take this factor into account, a second main research question is therefore posed:

2. Will the number and distribution of lexical dissonances differ between the intermediate and advanced students?

This research question can also be split into several, more specific research questions, including:

a) *Is there a difference in number of lexical dissonances found between intermediate and advanced students?*

b) *Is there a difference in distribution of lexical dissonance between routes?*

c) *Is there a difference in distribution of lexical dissonance between effects?*

d) *If there are differences between intermediate and advanced students, are these differences statistically significant?*

Taylor (1975) states that "with increased proficiency in the target language, [L2 speakers] rely proportionately less frequently on their native language grammar, and rely more frequently on their ever-increasing knowledge of the target language" (Taylor 1975: 88). This statement inspired a subquestion regarding L1-influence:

e) *Will the advanced students show proportionally less L1-influence in lexical dissonance than the intermediate students, illustrating "reliance on their ever-increasing knowledge of the target language" as suggested by Taylor (1975)?*

3. WORDS

Llach states that defining a lexical error is only possible once a word is defined (Llach 2011: 72). As words are not always as clearly defined as we think or wish, this section will clarify what is understood by the terms "word" and "lexical item," examine other common ways of classifying words, and look at some aspects of word knowledge.

3.1 Words and lexical items

There are several ways of classifying what constitutes a word. A word can be classified in terms of levels, such as the written, phonological and syntactic level (Saeed 2009: 56). Words can also be classified in terms of theoretical distinctions. Crystal suggests that there are three common theoretical distinctions of words, where words are defined based on if they are a) physically definable units divided by space, b) lexemes (variants of the same unit) or c) grammatical words (Crystal 2008: 522).

For this thesis it is more practical to define a word (and thus a lexical error) based on **meaning** rather than orthographic, syntactic or other features. For instance, the dissonant phrase: *As **late** as yesterday* (*as recently as yesterday*) illustrates that the student has chosen **one** wrong word, but it is only wrong because it does not fit in with the rest of the phrase. In other words, for a word to be wrong, it often has to be incompatible with the surrounding words, and these surrounding words can cluster together to form one unit of meaning.

Cruse explains that words can range from being the smallest mobile units in a sentence, to the largest units that "resist interruption" (Cruse 1986: 35-36). An uninterrupted unit can stretch from the single word *recently* to the phrase *as recently as* [time specification]. However, not many native English speakers would agree that *as recently as* is a single word. This is where the term "lexical item" comes in. Lexical items are generally viewed as one or several items which convey one meaning. This thesis counts such lexical items as one word or one dissonance – a "wrong word" can mean anything from a single word to an uninterrupted stretch of meaning which consists of several words.

Moreover, a common distinction is that of lexical words and grammatical words. Lexical words comprise main verbs, nouns, adjectives and adverbs, while grammar words consist mainly of determiners, pronouns, auxiliary verbs, prepositions and conjunctions (Hasselgård et al 2007: 14). As the focus of this study is on lexical errors, only lexical words will be analyzed.

Another common way of classifying words is if they belong to the same **word family**. A word family is a group of words that are grouped together on the basis of their morphology. They have a base form, possible inflectional forms and words that derive from them by prefixation and suffixation (Jackson & Amvela 2007: 19). For example, the word family PERFECT consists of the noun *perfection*, the verb *to perfect* (and all its subsequent verb forms), the adjectives *perfect* and *imperfect* and the adverbs *perfectly* and *imperfectly*. The term "word family" is useful for vocabulary acquisition researchers, because although a person may "know" 20, 000 words in a language, s/he may in reality only know, say, 5, 000 word families. This study uses the concept of word families to account for dissonance where the word base is correct (i.e. that it belongs to the right family, such as *perfect*), but the word form is incorrect (e.g. *a* ***perfectly*** *day*).

3.2 Word combinations and the relationship between words

> Words do not exist in isolation. Their meanings are defined through their relationships with other words and it is through understanding these relationships that we arrive at our understanding of words (Richards 1976: 81).

This thesis has mentioned "units of meaning," but not what these units of meaning consist of. For instance, in the case of *as* ***late*** *as yesterday*, it was determined that *late* is a wrong word choice. What constitutes these choices? This section will discuss "units of meaning" and how they relate to lexis.

3.2.1 Frequency and word combinations

One cannot talk about units of meaning, or indeed lexical choice, without discussing phraseology. Phraseology is the study of the structure, meaning and use of word combinations (Cowie 1994, quoted in Granger & Paquot: 27). Phraseology stresses that words are not lone units, and there are often patterns of co-occurrences – i.e. some words combine more often than others. Frequency is key to identifying word combinations – the more frequent two words co-occur (*blond hair*, but not ***blond*** *car**), the stronger the relationship between two words are said to be. Researchers believe that a speaker's ability to map and determine frequency and co-occurrence of words can only come through vast amounts of exposure to the

language (Schmitt 2010: 31). However, if words are taught in context (e.g. information about frequency of a word and its possible combinations are provided), it is believed that learners can use this information to quickly sound more native-like without being exposed to the language over a long period of time.

3.2.2 The illusion of choice

Sinclair, who is one of the advocates of phraseology, states that the role of words in sentences can be viewed in two ways: through the **open-choice principle** (few constraints on sentence/phrase structures other than grammatical ones which can be explained) and the **idiom principle** (semi-constructed phrases that constitute single choices, even though they might appear to be analyzable in segments) (Sinclair 1991: 110). The open-choice principle is problematic, because although there should be limitless ways of combining words that do not violate the rules of syntax, native speakers "do not exercise the creative potential of syntactic rules to anything like their full extent" (Pawley & Syder 1983: 193). This is referred to as the puzzle of native-like selection – how native speakers can distinguish between "which well-formed sentences are native-like" (Pawley & Syder 1983: 194).

The idiom principle can explain these native-like choices, by stating that these choices are determined outside the rules of grammar. There are many terms for these semi-constructed phrases including "formulaic language," "prefabricated language" and "phraseological units," to mention a few. The strength of how set this prefabricated language is can vary from words that are loosely grouped together, such as open slots (***a*** *good/great/bad/horrible* ***time***), to idioms that cannot substitute a single word (*all bark and no* ***chomp****).

Prefabricated language is no small part of language. A study conducted by Erman and Warren suggests that over 50% of a text will consist of prefabricated language, and for 100 words, only 71 word choices are possible (Erman & Warren 2000: 50). Aitchison postulates that multi-word expressions outnumber single words by ten to one (Aitchison 2012: 115). These types of prefabricated language are so vital to how speakers perceive lexis that there is a growing body of work indicating they are stored as automatized units in memory (i.e. that they are memorized and perceived as one "chunk") (Wood 2001: 579). The speed of retrieval of prefabricated language is so fast that this suggests prefabricated language is stored as a single unit in the mind. For native English speakers, the retrieval and combining of lexical chunks and frames is the foundation of fluency (Wood 2001: 582).

Although a speaker has hundreds of thousands of words at their disposal, the

restrictions on these words are immense and largely a result of co-occurrence frequencies. The ways in which words can combine can be utterly confusing to the non-native speaker, since there are no perceivable "rules" other than that some words tend to group more frequently with others. Followers of Sinclair and the idiom principle will stress the limitations of choice in lexis, and that these limitations are so imbedded in the language that they even affect how native speakers perceive and process language. L2 speakers sound "foreign" because they "exceed" these limitations.

3.2.3 Collocation

Strongly connected to the idea of the idiom principle is the concept of collocation, which can account for a large amount of lexical dissonance. **Collocation** is the frequent and habitual co-occurence of words in a language (Fiedler 2007: 189). It is divided into two types: **lexical collocation** (two lexical items: *strong tea*) and **grammatical collocation** (lexical word and grammatical word: *bring up*). Collocation consist of a node word (*tea*) and its collocate (*strong*). The node word and collocate are so inter-dependent that if two words which cannot collocate with each other are combined, dissonance is created between the node word and its collocate (***muscular/sturdy/tough*** *tea**). For the purpose of this thesis, lexical collocation is viewed as one lexical error. For instance, ***sturdy*** *tea* can be said to be one lexical item, and therefore will only be counted as one error.

Grammatical collocation consists mainly of multi-word verbs and delexical verbs. **Multi-word verbs** are verbs where there is a main verb and one or two particles that are viewed as constituents of a single unit (Jackson & Amvela 2007: 75). Some multi-word combinations are: prepositional verbs (*apply for, laugh at*), phrasal verbs (*bring up, count* [someone] *in*), phrasal-prepositional verbs (*look up to, catch up on*) and combinations that include three or more constituents (*make do with, take care of*) (Hasselgård et al 2007: 154-155). **Delexical verbs** are (main) verbs used so frequently that the meaning of the verb becomes reduced and is more dependent on the patterning of the words around it (Hasselgren 1993: 26). Hasselgård et al. list the five most common delexical verbs in English: *give, have, make, take* and *do*. Examples include: ***take*** *a look around,* ***have*** *a chat* (Hasselgård et al 2007: 156-7). Prepositions are disregarded in this study, as they are grammatical words, so prepositions in lexical items such as in *bring* ***in*** *a topic* are not counted. Only the lexical word can be counted in grammatical collocation (e.g. ***take*** *up a discussion*, instead of *bring up*).

Collocation links words together independent of sense. Aitchison posits: "Word

meaning is probably learned by noting the words which come alongside" (Aitchison 2012: 114). In other words, although there is no direct sense link between collocates, a reason for why there are such strong links between seemingly unrelated words could be because we create meaning **between** words based on how they are grouped together, independent of semantic fields. Lexical dissonance is termed "dissonance" because a large number of lexical errors are due to the fact that non-native speakers cannot see the link between these non-sense-related words – the link is only apparent through repeated exposure to the word combinations.

3.2.4 Idioms

Idioms are also regarded as one unit of meaning in this study of lexical error. There are several definitions of idioms. Jackson and Amvela explain that they can be viewed as: "a type of collocation involving two or more words in context [or] a type of multi-word lexeme," but that in general terms, an idiom is "defined as a phrase, the meaning of which cannot be predicted from the individual meanings of the morpheme it comprises" (Jackson & Amvela 2007: 77). This study counts both definitions as being a "lexical unit." It was found necessary to include them because wrong word choices can occur in large stretches of units of meaning. For instance, the phrase *want nothing to do with* is marked as stylistically dissonant in one of the student texts. It is impractical to label this as 3 lexical errors (*want, nothing* and *do*), since it is clearly the entire unit of meaning that has been misunderstood. Therefore, idioms, or longer stretches of units of meaning, have to be taken into account in this analysis of lexical errors.

3.3 Word knowledge

What does it mean to "know" a word? Or, how much information can a speaker deduce from a word? These are questions that many researchers have attempted to answer, particularly in relation to second language acquisition and learning, where learning words is a much more active process.

Jack Richards (1976) is the first to explicitly express that each word has several aspects of knowledge, and only by knowing these aspects, can we fully "know" a word. He lists what he believes these aspects of word knowledge are, with the purpose of offering a frame of reference for assessing vocabulary teaching. He examines linguistic, psycholinguistic, and sociolinguistic aspects of word knowledge in his article, including word

frequency, vocabulary growth in native speakers, collocation, register restraints, case relations, underlying forms, word association, and semantic structure (Richards 1976: 77). He famously formulates them as seven word knowledge assumptions, and these are regarded as the basis of the mapping of word knowledge (Richards 1976: 79-83).

Paul Nation, one of the biggest names in vocabulary acquisition research, refines and builds on Richards' word knowledge assumptions, listing the following aspects of word knowledge: **meaning, written form, spoken form, grammatical characteristics, collocation, register constraints, frequency** and **associations** (Nation 1990, quoted in Schmitt 2010: 31). Nation also includes other features such as orthographic and phonological, which are perhaps the most common measurements when people assess their own knowledge of a word (as in, "can I spell and pronounce the word?").

3.3.1 How does knowledge of words progress?

The acquisition of some aspects of word knowledge is not straightforward. Contextually-dependent aspects of word knowledge, such as collocation, meaning and associations can contain a lot of information. For instance, one word can have several meanings depending on context, and the more meanings a speaker is aware of, the more knowledge the speaker has of the word. Henriksen (1999) suggests that word knowledge is not a question of known or unknown – it is a continuum from zero to partial to precise knowledge of a word. Schmitt, using Nation's suggested eight aspects of word knowledge, creates a continuum for each aspect, proposing how learners of English progress along this continuum (Schmitt 2010: 37):

Figure 1: Progress of aspects of word knowledge in the early stage

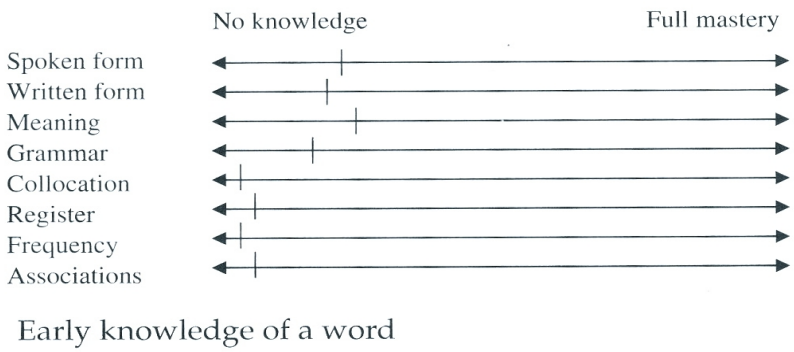


Figure 2: Progress of aspects of word knowledge in the developing stages

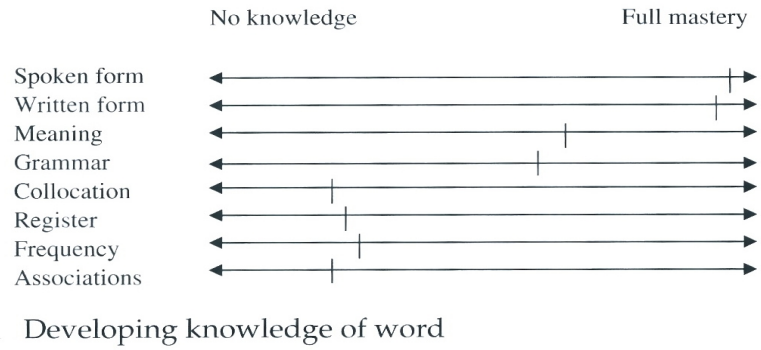
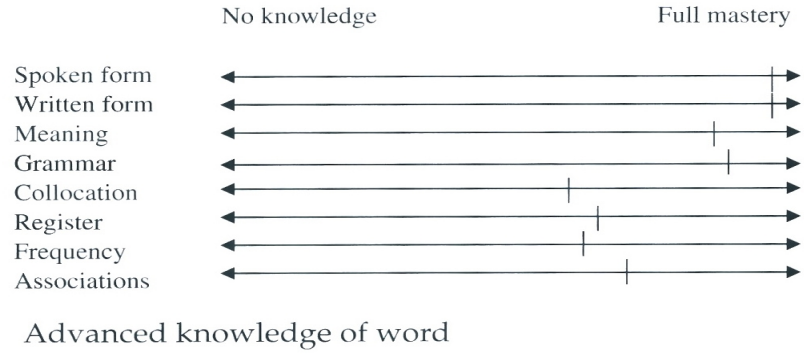


Figure 3: Progress of aspects of word knowledge in the advanced stages



In Schmitt's continuum of word knowledge, spoken form, written form and grammar are mastered first, perhaps indicating that these types of knowledge are limited and can be more easily mastered. Collocation, register, frequency and associations lag behind at all stages. Schmitt explains that these four categories are much more difficult to acquire because they do not occur in isolation and can therefore only be acquired through large amounts of exposure to English (Schmitt 2010: 31).

This continuum of word knowledge is relevant to this thesis because it examines "incorrect word hypotheses" (labeled "effects." See section 5.2.2). An incorrect hypothesis is the way a learner assumes something incorrectly about a word. These incorrect hypotheses can be incorrect at the level of "meaning," (labeled *semantic*) "collocation," "register," (labeled *style*), "grammar" (labeled *syntax*). They occur mainly when the learner only has partial knowledge of a word (i.e. early or developing knowledge of a word). An abundance of incorrect hypotheses in one aspect of word knowledge (for instance noticeably more collocational errors than other types of error) may indicate that the particular aspect is more difficult to acquire. This will not be examined in this study, but is something to keep in mind during the data analysis.

4. ERRORS AND THEIR SOURCES

An error is an unacceptable form of language found in spontaneous speaking or writing (Crystal 2008: 173). Errors occur if a speaker only has partial knowledge of a language, and are assumed to reflect a speaker's level of competence in a systematic way (Crystal 2008: 173). Error is defined by (un)intention – an error "only arises when there is no intention to commit one" (James 1998: 77). If a linguistic element is used in a different (perceivably wrong) way intentionally, it is not erroneous, but *deviant*. This is problematic for corpus data, because errors can only be deduced through context – it is unknown whether a learner intentionally deviates or not.

There are several degrees of error. James distinguishes between **slips** (lapses that are quickly detected), **mistakes** (can only be corrected if pointed out by someone else), **errors** (cannot be self-corrected until input has been provided) and **solecisms** (breaches of rules that even native speakers will make, e.g. split infinitives) (James 1998: 83). Mistakes and errors are of primary interest in error analysis, since slips do not necessarily reflect a misunderstanding in word knowledge, and solecisms are not a distinct L2 trait. Corpus data cannot easily distinguish between mistakes and errors, because there is no insight into the learner's mind. Therefore, all unacceptable, non-native forms (slips, mistakes and errors) can only be labeled as "errors."

This section will explain what is meant by "lexical error" and examine some of the sources of errors.

4.1 Lexical errors

Defining exactly what a lexical error is is problematic and complex. Generally "lexical error" refers to "the deviations in the learner's production of the L2 norm with regards to the use in production and reception of lexical items" (Llach 2011: 71). The line between grammar and lexis is fuzzy, so agreeing on what a lexical error is in relation to a grammatical error is is not an easy task. Some linguists interpret "lexical error" as meaning "all errors that are not grammatically fit" (Llach 2011: 73). Others view "lexical error" as a superordinate term for classes of errors such as word formation, collocation, form/semantic confusion and wrong word choice (Llach 2011: 73). Because of disagreement on what lexical errors entail, a large range of words is used to describe lexical error based on different definitions of what lexical errors represent. Examples include: "wrong lexical choice," "lexical deviances," "vocabulary

errors," "incongruencies in lexical gridding," "semantic deviation" "lexical confusions" and "lexical simplification," to mention a few (Llach 2011: 74). Despite the fact that researchers are reluctant to agree on exactly what lexical errors are and how they can be classified, there is a general consensus that lexical errors are a response to systematic causes and can be explained, classified and generalized (Llach 2011: 74). Research on lexical errors is often devoted to defining what a lexical error is to the particular study, creating a comprehensive taxonomy and then unearthing the systematic causes of lexical error for the data in question. As Llach puts, it "particular lexical error types are named depending on the dimension of the lexical error that prevails in the research," i.e. the definition and role of lexical errors is dependent on the data, and often taxonomies are built based on these needs (Llach 2011: 74).

In this study, a lexical error is a lexical item that deviates from L1-use. The lexical errors are referred to as "dissonances." One can say that "lexical error" and "dissonance" can be used interchangeably, but the term "dissonance" is preferred in this thesis. The term "error" "carries overtones of downright wrongness of meaning," hence the lengths researchers will go to to find other terms to describe this phenomenon (Hasselgren 1994: 238). A failure in form-meaning linkage (e.g. thinking that the word *cup* represents a vase) is perhaps the only way a learner's perception of a word can be downright wrong. This is because the speaker has failed to understand either the reference of the word (what a cup represents in the real world) or the form of the word (that it is called a vase). Most other instances of lexical errors are caused by a wrongly chosen word creating "*dissonance* between the word and its context" (Martin 1984: 130). There is often nothing wrong with the actual sense or understanding of individual words in "wrong word choices," such as the collocational mistake *nerves of* ***iron***. Context is key in detecting lexical errors – even instances of form-meaning linkage can only be located through mismatch in context in free production texts, e.g. *She dropped her phone on the* ***ceiling*** instead of *floor*.

4.2 Sources of errors

Richards and several other linguists have observed that non-native speaker errors derive from two major sources: interlingual and intralingual influence. In my taxonomy of lexical errors, source of error is one of the aspects which will be investigated. In this section, an account of these two sources of errors will be given.

4.2.1 Interlingual influence

Interlingual influence is influence which occurs between two or more languages. In practice, this normally means from a speaker's first language to their second or third language (or however many languages a speaker learns after their native language), but in many instances this is not the case. Observations from language acquisition studies have shown that all previously learned languages can influence additional language learning (i.e., L1 and L2 can both influence the L3 by facilitating or hindering L3 learning), and evidence is growing for cases of **bidirectional transfer**, in which influence can work both ways, from L1-L2 and L2-L1 (Ortega 2009: 48; 50). Therefore, many linguists prefer the term "cross-linguistic influence," because it is:

> theory-neutral, allowing one to subsume under one heading such phenomena as *transfer, interference, avoidance, borrowing* and L2- related aspects of language loss and thus permitting discussion of the similarities and differences between these phenomena (Sharwood & Kellerman 1986: 1, as cited in Ellis: 351).

This thesis uses the terms "L1-influence" and "interlingual influence." This is because although there is of course a strong case for cross-linguistic influence, the participants in this study only speak two languages (English and Norwegian), and only one type of interlingual influence will be examined – the influence from their L1 (Norwegian) to their L2 (English).

There are many ways of identifying L1-influence in L2 errors. This thesis differentiates between two types of influence: strong influence and weak influence. In lexical studies, strong influence occurs when an L2 speaker chooses a word or structure that is wrong in their L2, but right in their L1 (e.g. *he is **high*** instead of *tall* (*høy*)). The assumption is that the L2 speaker makes the lexical choice because of interference from their first language. This study also takes weak L1-influence into account, which for this thesis means words that are wrongly chosen because of the way some words behave differently in the L1 and L2. It is not a particular L1 word that is the cause of error, but rather that for example Norwegians who choose the word *human* instead of *person*, as in *she's a nice **human***, will have an interference, because Norwegian only has one superordinate word for the two English words (*menneske*). Sections 4.2.1.1 and 4.2.1.2 will explain some of the strong and weak L1-influences in detail.

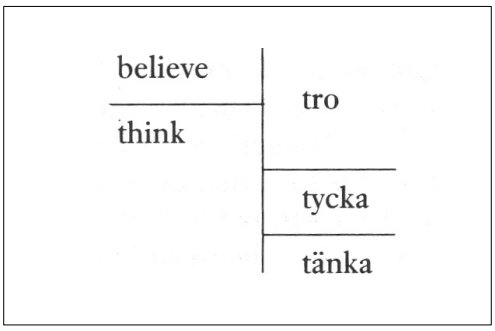
4.2.1.1 Strong L1-influence

When influence from one language causes a speaker to create a similar structure in the other language, it is often referred to as "**transfer**," i.e. that the speaker transfers their knowledge from one language to the other. Transfer is an important part of second language vocabulary

acquisition, as "generally the cross-linguistic influence between non-native languages in a European context has been shown to occur primarily in lexis" (Ringbom 1987: 114). There are two major types of lexical transfer from the L1 to the L2 that will be examined in this study: features of an L1 item that are assigned to an L2 item (lexical transfer), and L2 items which are combined in compounds or phrases analogously with the L1 structure (loan transfer) (Ringbom 1983: 207, quoted in Hasselgren 1993: 13)

Semantic lexical transfer is the extension of a word in the L1 to cover the semantic space of more than one L2 item. This can be explained in part by **the semantic equivalence hypothesis.** When two words are approximately equivalent to each other in another language, such as *think* and *tro*, Ijaz (1986) postulates that learners will rely on a "semantic equivalence hypothesis" (cited in Hasselgren 1993: 15). The learner will assume the semantic boundaries of an L2 item are identical to the L1 equivalent. Words cover different semantic spaces, and although a word in the L1 and the L2 may be translatable on some level, often the semantic spaces these two words cover will differ between the two languages. Take *think* and *tro* (Swedish), which have approximately the same main sense and can often be used in the same phrases and collocates in both English and Swedish. Then see how the semantic space is divided for the two:

Figure 4: Division of semantic space for English *believe/think* and Swedish *tro/tycka/tänka* (Carter 2012: 20)



A Swedish speaker of English who tries to directly translate *tro* to *think* may discover that her/his use of *think* is often wrong, since *tro* can also mean a personal belief or opinion, whereas *think* cannot. Conversely, English speakers of Swedish are often utterly confused when having to differentiate between *tro, tycka* and *tänka*, because they only have one word for these concepts. When semantic spaces do not correspond between two languages, as in the case of *think/tro*, the tendency is for L2 speakers to overextend the semantic boundaries of their L2 to "match" their L1. A comparison of semantic boundaries between L1 and L2 items shows that semantic boundaries which are similar are **convergent**, while semantic boundaries

that are dissimilar (*think/tro*) are **divergent**. The more similar two languages are, the more convergent these semantic boundaries will be. Divergence is nearly always caused by moving from a "coarser" to a finer "grid," e.g. moving from the L1 item *tro* to *think/believe* (Hasselgren 1993: 54). Such types of semantic lexical transfer are known as **finer gridding** – when the transition from a coarser L1 (Norwegian) grid to a finer L2 (English) grid causes the wrong word to be chosen.

However, evidence from this study shows that even non-semantic aspects of words are transferred from the L1 to the L2. This is reflected by collocational transfer (collocating words based on how they combine in the L1), stylistic transfer (words that are perceived to have a 1:1 relationship between L1 and L2, and which are stylistically appropriate in the L1 and not the L2), syntactic transfer (tacking syntactic features of an L1 item on an L2 item) and the creation of non-existent phrases or words in the L2 that derive from the L1 (one student in my study wrote ***office rat*** (*kontorrotte*) to denote people who spend too much time at the office). Because of these instances found in my data, I would suggest that the semantic equivalence hypothesis be stretched to a **word equivalence hypothesis** – if an L2 speaker perceives and L1 and L2 word as having a 1:1 mapping, they will often transfer knowledge of their L1 to alter the semantic space of an L2 item, its word combinations, stylistic markings and even syntactic features.

The second major type of lexical transfer is **loan translation**. Loan translation is "the use of the literal translation of compound words, lexical collocations, or idioms from one language to another" (Ellis 2008: 370). This will be further explained in section 5.2.1.2.

4.2.1.2 Weak L1-influence

Finer gridding can be the result of lexical transfer (such as in the case of a Swede overusing *think*), but not necessarily so. For instance, how can semantic lexical transfer explain if a student overuses *believe* instead of *think*? Which of the L2 items will the L2 speaker pick as the "overextender?" In cases where L1 items are subdivided into two or more items, such as *tro*, which has components common to *think* but also additional meaning (*believe*), it has been observed that the learner will "intuitively apply **the most familiar** of the L2 items to the whole semantic area in question" (Hasselgren 1993: 15, emphasis mine). A Swedish speaker of English may therefore prefer *think* because it is more similar to *tro* in form and main sense, or s/he may prefer *believe* if this is the L2 item s/he has heard the most and is the most familiar with. In the case of *think/tro*, picking *think* indicates a strong L1 influence due to

semantic transfer, because the student (most likely) perceives a 1:1 mapping between a familiar item (*think*) and an L1 item (*tro*). In the case of *believe/tro*, *believe* indicates a weak L1-influence, because the problem lies in having to divide between semantic spaces that do not exist in the L1, rather than transferring semantic aspects of the L1 to the L2. The incorrectly picked word was not influenced by the L1. Instances of finer gridding where an L2 item that is not similar to the L1 is picked is therefore termed **weak finer gridding**, and labeled as weak L1-influence.

Another instance of weak L1-influence is the overuse of basic words (known as "core words" in this thesis; see section 5.2.1.6). Often basic English words will have a Norwegian equivalent (e.g. *big – stor*). When a Norwegian speaker of English chooses to overuse *big*, it is more likely preference for a basic word, rather than an L1-influenced choice of words. However, basic words in English and Norwegian behave differently. For instance, *big* is seen as much more informal, such as *big question*. It has an informal stylistic mark, so L1 English speakers will prefer synonyms that have a more formal mark when writing formal texts, such as *large*. Norwegian *stor* does not have an informal stylistic mark, so writing *stort spørsmål* is more acceptable in academic writing. Using basic words in academic writing shows clear dissonance in English, but not to the same degree in Norwegian. Therefore, (some) basic words have weak L1-influence: it is possible that Norwegian speakers of English find them more acceptable in English because they are acceptable to use in their L1. Instances where basic words have an equivalent L1 that does not have the same stylistic markings are considered to be weakly L1-influenced. As in the case of weak finer gridding, the problem lies in having to divide the language differently from the L1. These instances are not considered intralingual influence, because if the L1 item had had a finer grid, the particular L2 item may not have posed a problem.

4.2.2 Intralingual influence

Intralingual influence is interference or confusion within the target language itself. Intralingual influence is unrelated to the L1 and can either reflect unique errors of the individual speaker or the developmental universal processes that are a part of language acquisition in general (Ellis 2008: 53, Ortega 2009: 51).

In vocabulary acquisition/use, developmental processes are often categorized by level of meaning, e.g. from concrete to abstract senses of words. A common intralingual problem is overgeneralization. L1 children overgeneralize concepts, such as calling all animals with four

legs “doggy.” However, L2 speakers have most likely developed these concepts in their first language, and overgeneralization of concepts tends to be replaced by L1-influenced semantic lexical transfer. L2 speakers overgeneralize by overusing superordinate, approximate terms, in instances where a native speaker would have used a more specific term , for instance *to **do** a study*, instead of *conduct a study* (Hasselgren 1993: 8). Wrong L2 words that have no perceivable L1-influence – especially between two languages that have little vocabulary in common – tend to be viewed as a result of intralingual influence. What causes these intralingual errors is unclear. López explains them as being a consequence of “partial exposure to the target language” (López, date of publication unspecified). Hasselgren stresses the challenge of choice that L2 speakers are faced with when “selecting words from an often confusing and incomplete stock of L2 words in the interlanguage” and refers to this selection as “lexical coping” (Hasselgren 1993: 1). Linguists who study vocabulary acquisition, such as Hasselgren (1993, 1994), Schmitt (2010) and Martin (1984), all seem to be in consensus that picking the wrong word is the result of a) lack of word knowledge and b) coping with lexical choices when there is a lack of word knowledge. Sometimes these lexical choices will be influenced by the L1, intralingual factors, a combination of the two, or other factors that researchers cannot yet explain.

In this study, words that do not have perceivable L1-influence are considered a consequence of intralingual factors, which could be the result of overgeneralization, the picking of the wrong synonym (arbitrary guessing or confusion of word forms), teaching-induced errors (students are encouraged to use overuse certain terms) and lastly, idiosyncratic errors which are most likely the result of a conceptual misunderstanding of the word within the individual speaker's mental lexicon.

4.2.3 Errors – interlingual or intralingual?

One of the heated topics in SLA literature and research is how large a role interlingual influence – particularly L1-influence – plays in acquisition and use of the L2. The role of L1-influence was stressed in earlier times, such as the rise of contrastive analysis in the 1950s-1960s. Contrastive analysis stated that differences between the L1 and L2 could purely be explained through a systematic comparison of the two languages. Any differences between the two languages could predict possible problems for L2 learners. The **contrastive analysis hypothesis** stated the following: in cases of similarity between native language and target language, the learner could just transfer their native tongue habits. If this transfer was successful, this would then be **positive transfer**. In cases where the two languages differ,

negative transfer (a structure that exists in your native language but not in the target language) could be predicted. However, the L1 cannot **always** account for the mistakes L2 speakers make. Analyses began showing that a large proportion of errors could not be predicted by contrastive analyses (Simensen 1998: 91).

With stacking evidence against the deterministic contrastive analysis, a weaker version of contrastive analysis, where only **some** errors could be explained by comparing the two languages, was suggested. This focus on explaining the source of errors was the beginning of error analysis, and examined both interlingual and intralingual errors. In modern times, intralingual influence has been studied and is more widely accepted as a large factor in L2 errors. Nevertheless, the dichotomy between interlingual versus intralingual influence causing the most amount of confusion and errors in L2 is still widely discussed. For instance, Ellis writes that "A large number – and in some cases perhaps most – of the errors that learners produce are intralingual in origin rather than transfer," but concludes that "the proportion of transfer and interlingual errors varies in accordance with the task used to elicit samples of learner language" (Ellis 2008: 55).

Yet the divide between inter- and intralingual influence is not as clear-cut as presented above. Errors can have more than one source, as illustrated in Hasselgren's treatment of influence on choice of wrong words. She finds the need to create an L1/L2 category which covers cases where she sees that "both languages seem to exert a simultaneous influence" (Hasselgren 1993: 46). She provides an example of an L1/L2 combination in the case of when one student chooses *overlook* as an appropriate translation of *overse* (ignore). Here there is a dual factor of literal translation and a word association error (see section 5.2.1).

Since errors can have more than one source, and it is impossible to know for certain what the speaker is thinking when producing the errors, several linguists are skeptical of labeling sources. Flick (1979), for instance, comments that assigning sources to errors is "largely an arbitrary matter, subject to the individual biases and point of view of the researcher" (Flick 1979: 60, quoted in Ellis: 55). Nonetheless, this thesis aims to point out potential sources of errors.

5. TAXONOMIES OF LEXICAL ERROR

There are many taxonomies of lexical error in SLA literature. Ostensibly, these taxonomies

have little in common. Llach, in an attempt to characterize them, lists eight main classification criteria, (Llach 2011: 76-87):

1) **Distinction between form-oriented and content-oriented lexical errors** to explain if the lexical errors semantic or formal in nature.

2) **Descriptive criterion** includes pure descriptive considerations that focus on the surface form of the error without alluding to causes or sources. Examples: wrong lexical choice, omission, wrong order.

3) **Etiologic or psychological criterion** explains the source of lexical error. Lexical errors are classified relative to their cause – mental processes underlying the lexical error. Example: overgeneralization, semantic transfer and confusion of related words

4) **Origin of influence criterion** includes interlingual, intralingual and teaching-induced errors .

5) **Grammatical or linguistic criterion** has to do with the classification of lexical errors depending on the linguistic level on which the error occurs. Phonology, orthography, morphology, syntax.

6) **Word class criterion** examines which word classes are the most affected by errors.

7) **Product-/process-oriented taxonomies** take psychological processes that generate errors into account.

8) **Miscellaneous.** Taxonomies that combine several classification criteria in an endeavor to establish a complete taxonomy that collects as many lexical error types as possible.

Taxonomies are in most cases developed to fit the data – not vice versa. Therefore, it is difficult for lexical error taxonomies to be adapted for other data sets, such as learners with different L1s, or different types of data (collected through elicitation method, judgment tasks etc) (Llach 2011: 91). Due to these issues, it seemed most prudent for this study to use a taxonomy of lexical errors created by a researcher who had similar data. The first criterion was that it should be a taxonomy created specifically for the Scandinavian L1s (Norwegian, Swedish or Danish), since the languages are similar to each other (for comparison) and similar to English (for taxonomies that focus on similar L1-L2s). A major figure in lexical error research within the Nordic countries is Håkan Ringbom (1987, 1992, 2007), who investigates

L2 English lexical errors in Swedish- and Finnish-speaking Finns. However, Ringbom only takes into account interlingual influences, and thus intralingual influences are not explained. His taxonomy was therefore not ideal. Hasselgren (1993), on the other hand, created a taxonomy for Norwegian speakers of English that also includes intralingual influence. Her taxonomy is of Llach's "miscellanous" type because it features criteria such as content-oriented, psychological (routes), origin of influence (L1/L2 influences) and descriptive (effects). A few modifications were needed to fit the taxonomy to my data since her data consist of translation texts and mine consist of free production texts. How the taxonomy is altered will be addressed in 5.2. Hasselgren does not start from scratch – her taxonomy has origins in Marilyn Martin's taxonomy of dissonance. Therefore, the next section will examine Martin's and Hasselgren's taxonomies of lexical dissonance, accounting first for Martin's, and how Hasselgren built on Martin's model.

5.1 Martin's taxonomy of lexical dissonance

Marilyn Martin wrote an article in 1984, titled "The problem of synonyms." In it, she attempts to explain why some advanced speakers sound "foreign" in relation to vocabulary. Martin concludes that dissonance arises when a student has made an incorrect hypothesis about the wrongly chosen word, e.g. that an informal word can be used in formal contexts (Martin 1984: 130). According to Martin, dissonance occurs first and foremost because students are not provided with enough information of new words in the foreign language classroom, be it glossaries or "synonyms." For instance, a Norwegian student of English who is taught that *large* means "the same" as *big* may be in for a surprise when they want to tell someone about their *large sister* (storesøster). The wrongly chosen words are synonyms that share some properties with the correct word (*large – big*), but there are several aspects of words that make synonyms different from one another. When L2 speakers are led to believe that such synonyms are completely interchangeable, dissonance takes place. Martin proposes that there are four types of dissonance that interfere with encoding by advanced learners, and once these dissonances are identified, language teachers will more easily be able to pinpoint why a word is "wrong" and furthermore teach words within this framework to avoid further mistakes. The four types of dissonance are[4]:

4 These types of dissonances are known as "effects" in this study, and will be further explained in 5.2.2. Martin presents a highly detailed list explaining her taxonomy. As these categories are treated slightly differently in Hasselgren's taxonomy, these will not be explained in great detail, but can be found in Martin's article, pages 131-134

I) **Semantic** (the speaker has misunderstood the meaning of the word, e.g. superordinate/ hyponym)

II) **Stylistic** (the speaker has misunderstood the style of a word, e.g. formal /informal/ taboo)

III) **Collocational** (the speaker is unaware of combinations of prefabs)

IV) **Syntactic** (the speaker is unaware of the words that tend to surround the dissonant word, e.g. case relation confusion)

Martin suggests that if these four types of dissonances are explained in detail to the language learner, many synonymous errors can be avoided, and ideally L2 speakers can teach themselves these different aspects of word knowledge, to avoid habitually using words interchangeably.

Martin's taxonomy for lexical dissonance is a good starting point for any lexical researcher because it is a comprehensive taxonomy that can account for why several types of words can be dissonant. However, there are several shortcomings with her model, i.e. there are dissonances between meaning and context that fall outside the framework of her categories, and thus cannot be explained by them. Carter points out among other things that grammatical and lexical components combined will also result in dissonance, e.g. *He ran very fast* ***and*** *failed to win the race* (Carter 2012: 83). Additionally, dissonance does not pertain only to the clause-level. Carter illustrates this with beyond-the-sentence errors, such as: *She passed the exam.* ***This move*** *pleased her parents* (Carter 2012: 83). He explains this as:

> To account for and understand lexical patterning at the level of the clause, ways with words depend just as crucially on the patterns created by lexical items in the wider context of naturally occurring discourses (Carter 2012: 83).

Another problem with the taxonomy is the assumption that only wrongly chosen synonyms can account for lexical dissonance. Occasionally an L2 speaker will use a word without knowing it at all, or use a word and perceive it to mean something completely different. The end result is at best a word that is unrelated to the right word yet guessable through context, and at worst an entire clause/sentence where it is impossible to understand what the speaker means. An example of this is my definition of "perceived equivalents" (see 5.2.1.3), where L2 speakers will use words that cannot be explained by any of the suggested incorrect hypotheses in Martin's framework. For instance, one student writes: *A* ***division*** *in time*. This could possibly mean *gap* if we look at the surrounding words and our knowledge of words that

collocate with *time* – but *division* and *gap* are not synonyms. If *gap* is not what the student meant to express, I have no further suggestions. Wrong words in this category are found even at the advanced level of Norwegian speakers. The reason for why unrelated words can be picked instead of synonyms is that failure to encode the word (e.g. a form-linkage failure) can cause students to arbitrarily pick a word they do not know at all, instead of simply using "similar" words interchangeably.

Martin's neglect of L1-influence also results in a lack of coverage of dissonant items. L2 speakers are almost referred to in her article as less knowledgable L1 speakers, rather than second language learners who have an abundance of knowledge and resources to call upon when they are in doubt. It is impossible to claim that L2 speakers only refer to the second language when speaking/writing in the second language – they often draw on knowledge of their native tongue when they are unsure of a new word, and create "incorrect hypotheses" based on their L1 knowledge, e.g. semantic lexical transfer. Seasoned second language teachers such as Hasselgren recognize the importance of the L1 – even in advanced speakers of English. The semantic equivalence hypothesis is an example of this – when some aspects of words are unknown, L2 speakers rely on the structure of their own language (*high* - *høy*), rather than transferring knowledge of other, similar L2 synonyms (*high – tall*).

5.2 A taxonomy of lexical dissonance for Norwegian students of English

There are several weaknesses with Martin's taxonomy that need to be addressed if it is to be practically implemented. First, it is purely descriptive in nature because it does not explain the sources or mental processes behind choosing the wrong word. Another shortcoming is that L1-influenced dissonances which may create items that do not exist in the L2 (e.g. *office rat*) cannot be explained by the taxonomy. In Hasselgren's examination of Norwegian student texts, these shortcomings needed to be addressed in order to answer the following questions:

i. **Why** do learners choose wrong words? (Influence)

ii. **How** do they choose wrong words? (Route)

iii. **What** is wrong with the choice of words? (Effect)

Such an investigation needs among other things “origin of influence” and “psychological” criteria to explain the how and why in the choice of wrong words. Therefore, Hasselgren creates her own taxonomy for “routes” (how learners choose wrong words, or psychological criterion) and “divergence” (why do learners choose wrong words, or origin of influence), in addition to Martin's taxonomy, which she labels “effects” (what is wrong with the choice of words). The process from choosing a wrong word to its effect can be explained in the following way by Hasselgren's taxonomy: a learner does not know how to express *tall*, as in *he is tall* (*han er høy*). The learner knows Norwegian *høy*, which has an English cognate, *high*. The learner also knows that *high* is an English word, and assumes a semantic equivalence, i.e. that *høy* can be directly translated to *high*. The learner then chooses this cognate as a “route,” or pathway, and the learner chooses the wrong word because s/he relies on her/his knowledge of L1 (Norwegian). The belief that *høy* = *high* is an incorrect hypothesis, because *high* is not used to describe people in English. The effect of the wrong word is semantic dissonance.

According to Hasselgren's taxonomy, Norwegian speakers of English use six routes (pathways that can potentially lead to dissonance), and there are five different ways in which a word is wrong (effects). The dissonances can be caused by L1-influence, L2-influence, L1/L2 influence or other influences (e.g. obvious avoidance of certain terms). The next sections will outline routes, effects and influences in detail. Since Hasselgren uses translation texts and I use free production texts, I found the need to modify and add to some of her categories.

5.2.1 Routes

What is a route?

A route is *how* learners make mistakes, or *in what way* a learner selects a wrong word (Hasselgren 1993: 45). Hasselgren created the taxonomy “routes” herself by examining large bodies of texts written by Norwegian speakers of English and determining the different ways students select wrong words. She posits that there are six ways in which Norwegian students of English misselect words: cognates, transliterations, perceived equivalents, synonyms, associations and cores. Some of these concepts are by no means new: Ringbom has been investigating the role of cognates and transliterations in learner language since the early 80s. However, Hasselgren is the first to build a system that could explain all possible ways Norwegian speakers of English misselect words.

5.2.1.1 Cognates

A **cognate** is "a linguistic form which is historically derived from the same source as another language/form" (Crystal 2008: 83). An example of this is English *see*, which has a Norwegian (historically derived) equivalent: *se*. In many cases, cognates lead to positive transfer – they are easier to acquire and use, such as in Ringbom's studies of Finnish versus Swedish speakers of English (1987, 1992, 2007), where the Swedes have a great advantage over the Finnish. He explains this as: "when both phonological and semantic similarity work together, the effect is like that of a magnet attracting a new word to be stored in the learner's mental lexicon when he meets it for the first time" (Ringbom 1987: 41). In other cases, they are misleading, such as when they cover different semantic spaces or collocate differently. Equivalent cognates lead to positive transfer and facilitate second language learning, whereas non-equivalent cognates lead to negative transfer. This is stressed by Ringbom, who discovered that **false friends** (a Swedish and English word that have formal similarities) were the highest source of errors due to Swedish influence in his study (Ringbom 1987: 124).

Cognates always have L1-influence, which means that they are L1-based. A wrong word in this study is labeled a cognate if the L2 item has an equivalent L1 item that is similar in **form** and (occasionally) **meaning**. Examples of this are: *train – trene* (*I **train** at the gym*), *mean – mene* (*I **mean** that fruit is healthy*), *clock -klokke* (*What is the **clock**?*) and *land - land* (*Norway is a nice **land***). If the word is historically derived from the same source or not is not a criterion (although in many cases it will be). My reasoning is that when a Norwegian speaker of English decides to perceive two words as being the same in both languages, they do not consider whether the words are linked together *historically*, but rather if the forms and meanings are close enough for there to be a 1:1 relationship between the two. Words that are similar in meaning but **not** in form are considered transliterations (for instance, *live - bo*).

5.2.1.2 Transliterations

Transliterations, much like cognates, stem from L1-influence. Hasselgren explains transliteration as: "the process of breaking down an L1 item, and reassembling it with L2 parts, often resulting in a non-existent word or phrase" (Hasselgren 1993: 48). Sometimes breaking down an L1 item will result in positive transfer – particularly if the two languages in question are similar. Any instance where you correctly guess a word or expression in a foreign language with L1 influence is an illustration of successful transliteration. In cases where L2 knowledge is lacking and a guess based on L1 knowledge is made unsuccessfully, dissonance

caused by transliteration occurs. The main difference between cognates and transliteration is that if the **form** of the two words is similar, they are classified as cognates in this study, and if they are directly translatable words that are **not** similar in form (*fin - nice*), they are classified as transliterations. Examples of transliteration are: to ***live*** *at the hotel for the weekend* (*bo på hotell*), to ***make*** *a party* (*lage en fest*), and to ***do*** *a good* ***try*** (*gjøre et godt forsøk*).

5.2.1.3 Perceived Equivalents

Perceived equivalents are defined by Hasselgren as:

> a small but widely used group of L2 words felt instinctively to be chosen under the influence of particular L1 items, yet having no formal resemblance to them. An example is the verb *experience*, frequently provided when *find out* or *discover* would be more appropriate (Hasselgren 1994: 241).

Perceived equivalents in this analysis are treated differently. The reason for this is that Hasselgren had access to Norwegian source texts. In free production texts, there is no such insight into what the student was supposed to or wanted to express. It is therefore impossible to determine what the student believes to be a 1:1 mapping, unless the student repeats the same mistake several times in different contexts, showing that when they write *experience*, they clearly mean *discover*. This analysis includes such repeated errors where context shows a false 1:1 mapping between the Norwegian L1 item and the English written item. Such repeated errors do not have any perceivable L1-influence in my study, and are often explained by unsuccessful coding of the L2 word, or other unknown factors. The criterion for a perceived equivalent in this study is a wrong word that is repeated several times, and where there is no perceivable L1-influence. An example of this is when one student uses the word *profit* instead of *goal/aim* twice in a text. Example: *Maybe the* ***profit*** *is to reach popularity. Profit* and *goal/aim* are not related in any way, either semantically or by word form. There is no perceivable influence from the L1 because Norwegian differentiates between *profit* (profitt/overskudd) and *goal/aim* (*mål*), and neither of these L1 items should be confusing to the L2. The student also repeated this error 2 times, indicating that s/he **perceived** *profit* to mean *goal*.

Perceived equivalents also includes words where the intended meaning is simply unknown. Examples of unknown intended meaning are: *she gave him a* ***withering*** *look*, or *a* ***mortal*** *glance*. The reason why these unknown dissonances are considered to be perceived equivalents is because there is an underlying assumption that when an L2 speaker produces a sentence, they have a clear picture of what they intend to say (in other words, they do not just

write words arbitrarily without some intended meaning they wish to convey), whether they have a Norwegian L1 item in mind or not. So while the student perceives the word they write to have the meaning they wish to express, either the word form or sense is wrong in some way, and the intended meaning cannot be determined (unless it becomes clear through context).

Since this route does not have perceivable L1-influence in my data, it is labeled as L2-based. The criterion for an L2-based route is that the difficulty lies in coping with the choices and restrictions of the L2, rather than interference from the L1 (Hasselgren 1993: 50). However, as in the case of perceived equivalents, problems could also mainly lie in the encoding of information of the word.

5.2.1.4 Synonyms

Synonyms are words with approximately the same meaning. However, no two words can mean *exactly* the same. The fact that words can be seemingly similar yet differ on many levels is a point stressed by Martin, who created the whole category "effect" to explain why choosing a synonym can mean choosing the wrong word.

Categorizing the route "synonyms" is tricky , for are not most dissonances the result of simply picking the wrong synonym? "Synonyms" is therefore a miscellaneous category for dissonant words where the intended meaning is clear, but there is a) no direct L1-influence (cognates and transliterations), b) the wrong word and correct word are not similar in form[5] (e.g. ***take*** *an assumption* instead of <u>*make*</u> *an assumption*) and c) the synonym is not an overused basic word where several other correct items would have been appropriate[6] (e.g. *nice* instead of *wonderful, splendid, fantastic*). Synonyms belong to the same word class and often the same semantic or associational field (i.e. the words are related in meaning or grouped together by how we associate them, e.g. *pillow – bed*).

"Synonyms" is L2-based because the problem does not lie in direct transfer of word knowledge from the L1 to the L2. Synonyms can also have L1-influence, but this influence is weak, as in the case of weak finer gridding (see section 4.2.1.2). Unlike the L1-based cognates and transliterations, many synonymous mistakes bear no resemblance to Norwegian translations/equivalents. Examples of synonyms: *the volcano* ***detonated*** (*erupted*), or *a smile*

5 These dissonances are labeled "associational" dissonances

6 These dissonances are labeled "cores"

*crossed her **mouth*** (*lips*).

5.2.1.5 Associations

Associational dissonance is dissonance where a) the L2 speaker has chosen an incorrect synonym that is similar in form (and possibly meaning) to the correct word or b) the student has spelled a word so incorrectly it somehow changes the form of the word (such as *sward** instead of *swore*). Association bases itself on the dual factors of form and (loose) meaning association, i.e. words that sound alike or are written similarly are confused with one another when selecting a synonym (Hasselgren 1994: 242). This confusion of form and (possible) meaning association does not only pertain to L2 speakers. Learner errors caused by association with phonologically related words (*take - make*) are strikingly similar to native speaker "slips" (Meara 2009: 28). Selection errors (a type of "tongue slip" where the wrong item is selected from the mental lexicon) can be based on **meaning similarity** (*He came **tomorrow*** (*yesterday*)), **sound meaning** (*The emperor had several **porcupines*** (*concubines*)) or **meaning and sound** (*I don't have empathy for rich-looking **burglars*** (*beggars*)) (Aitchison 2012: 22). In this study, associational dissonance is based on sound similarities or a combination of meaning and sound. Sound similarities are called **malapropisms** – cases in which a similar-sounding word has been wrongly selected, as in *cylinders* for *syllables* (Aitchison 2012: 158-159). Words that have a meaning similarity (***tomorrow -*** *yesterday*) are placed in the synonym category.

Associational dissonance also takes the relationship of lexemes (pertaining to word families) into account, such as to *hold a **speak*** (*speech*), or *her hair looked **beautifully*** (*beautiful*). If the wrong word class for the correct lexical item is chosen (*speak - speech*), these are labeled associational dissonance, since it could be the word form that is confusing.

It is difficult to determine where spelling mistakes end and associational dissonance begins. Accidental or common spelling mistakes (such as "**slowe**" for "slow") should be disregarded in lexical studies. Only dissonance that reflects a lack of knowledge of word sense, collocation, syntax or severe misspelling of word forms are considered. Several spelling mistakes which can be found in the texts are therefore omitted from the list of dissonances. Examples of associational dissonance include *an **exaggerated** birthday party* (*extravagant*), *to **increase** your looks* (*enhance*) and *taught* instead of *thought*. Associational dissonance is an L2-based route, because the wrong L2 item is not influenced directly by an L1 item.

5.2.1.6 Cores and Lexical Teddy Bears

Cores and lexical teddy bears are words that block the production of more appropriate words or a variety of synonyms. Cores and lexical teddy bears are not exactly the same, but both have the same effect of blocking production and often overlap. This section will explain what core words and lexical teddy bears are, why they are dissonant and how they are classified.

5.2.1.6.1 What is a core word?

As per definition, a **core word** is a neutral word with many senses that can collocate freely and is neither too formal or informal. Carter furthermore defines core words as: "central members of their lexical sets or superordinates, and they will collocate widely, frequently being extended into phrases" (as cited in Hasselgren 1993: 33). Core words often stem from **core vocabulary** – a group of widely used words that are more resistant to loss or change than other parts of vocabulary (Saeed 2009: 77). Normally, core vocabulary is established and based on frequency studies, particularly with the aid of corpora in recent times. Mapping core vocabularies is thought useful for two main reasons: 1) establishing how similar or different languages are and 2) to boost speed in L2 acquisition. Core words are considered ideal for quickly improving second language vocabulary, as they have many notions which can be re-expressed, and are often considered "basic" language (Carter 2012: 40). One project which attempted to map core words for such a purpose was the Basic English project, which sough to provide a basic minimum vocabulary involving a list of 850 words (Carter 2012: 37). However, projects such as the Basic English project often fall short, for even though it seems simple enough to learn 850 words in order to speak basic English, Nation points out that these 850 words can have as many as 12, 425 meanings (Nation 1983: 11, quoted in Carter 2012: 40), and research has not established whether it is easier to learn the extended senses of such core words, or to learn the different meanings of an almost equal number of lexical items (Carter 2012: 41).

5.2.1.6.2 Core words – wrong?

Even though core words are often characteristic of L2 learners and considered dissonant in some cases, they are by no means "wrong" to use. Native speakers make frequent use of core words, which is why they are considered "core" in the first place. A core word is only considered dissonant when a student uses a core word when a more specific, register-appropriate word would have suited better. For instance, the adverb *pretty* and adjective *good*

are perfectly appropriate in colloquial English, such as ***pretty good*** *results.* However, in an academic setting they are not register-appropriate, and should be replaced with e.g. *quite positive results*, or other synonyms that reflect the style of register. The reason why core words are not labeled as synonymous dissonance (which, of course, they are to some extent) is because core words reflect an interference from words learned early in the L2 stage. For instance, an L2 speaker who learns *good* at an early stage will be more reluctant to use or learn other, more specific synonyms, such as *admirable, satisfactory*, or *splendid* if their use of *good* feels adequate and is positively reinforced. Hasselgren explains this as:

> words with highly specific meaning and restrictions on use are unpopular - *got* will usually be preferred to *acquired* […] Multiplicity of meaning has been shown to lead learners to avoid using a word in more than one sense, a point reinforced by Laufer [...], who cites polysemy as the largest single source of comprehension errors (Hasselgren 1993: 19).

Overuse of core words is often characteristic of non-native speakers, as they lack some of the lexical nuance and density that a native speaker is more likely to possess. This is evident in Hasselgren's own study of "different words," where she concludes that:

> learners have a distinct preference for "core items." While native speakers tended to choose intensifiers that somehow capture the specific semantic area, Norwegians on the whole did not do this. When native speakers favor unrestricted items, such as *great* or *pure*, the learners tended to follow suit. However, when the favored native item was restricted, such as *profusely* or *utterly*, the two groups parted company (Hasselgren 1994: 254).

5.2.1.6.3 Labeling core words

Core words, particularly stylistic ones, are difficult to label. The reason is not because it is difficult to determine whether a core word is a core word or not – the difficulty lies in determining whether these core words are **inappropriate** or not in context, and if another lexical item would have been used by a native speaker. Some of the intermediate free production texts are literary. Core words are largely not considered dissonant in this genre, because the register is appropriately informal and colloquial. Even in academic writing there is a gray area between when a core word is considered inappropriate or not in context. For instance, *very* is generally considered inappropriate and is avoided in essays, such as in the example: *a* ***very*** *interesting article.* Most native speakers are aware of the redundancy of using this adverb, and will either avoid it, or use a stylistically appropriate synonym, such as *fascinating*. However, this does not mean that *very* is **always** dissonant in argumentative essays. In another instance, a student wrote: *these are shown in* ***very*** *different ways.* In this case, *very* is not redundant and adds to the text. When the core word list states that there are

38 instances of *very* as a core dissonance for intermediate students, it does not mean that the word *very* is only used 38 times – each context has to be taken into account to establish if the core word shows traces of dissonance or is used appropriately. So if a core word struck me as being inappropriate *in context*, it was labeled as dissonant.

5.2.1.6.4 Lexical teddy bears

The term "lexical teddy bear" has been used without much explanation. First coined by Angela Hasselgren, it is explained as: "stripped of the confidence and ease we take for granted in our first language flow, we regularly clutch for the words we feel safe with: our '**lexical teddy bears**' (Hasselgren 1994: 237, emphasis mine). This sense of "clutching" to words is considered a type of lexical dissonance because a word will become much more frequent than in native speaker use. The words we cling to in our second or third language are more often than not core words learned at an early stage, but not necessarily so. Occasionally we simply find words that we like, feel comfortable with and overuse. An example could be overuse of the word *really* (***really*** *good, r**really** great*), which does not have the neutral status of a core word (it is stylistically marked as colloquial), yet is overused. The danger of lexical teddy bears is that they may lead to imprecision (*he never managed to **take** the exam*), wrong collocations (*this didn't **awake** any interest*) which results in production blocking of other, more accurate words (Hasselgård et al 2007: 66). An example of a core word used as a lexical teddy bear: *The histories [stories] tell us a very **good** story […] I would say that if you live with a thread it's hard to live a **good** life […] I think that is **good** that some people write about this type of stories*.[7] Here, a multitude of synonyms could have been used, but the student clutches to the core word *good* to describe a wide range of senses. Although core words can be and often are overused, this is not necessarily the case. For instance, the core word *bad* is only found twice in the intermediate student texts. Since *bad* is stylistically incorrect, it is considered a core dissonance. However, since it is used so infrequently, it is not a lexical teddy bear. Setting a definite limit for when a word is a lexical teddy bear is a difficult task – it would be easier to put a word on a continuum where it is either frequently or infrequently used compared to L1 use. However, a limit had to be set for the purpose of this study. A lexical teddy bear is therefore a stylistically inappropriate word used over five times in a corpus (25 000 words), with or without being a core word.

7 Example taken from one of the students in the intermediate corpus

5.2.2 Effects

What are effects?

Once it is established *how* learners make mistakes, another question arises: *why* is it wrong? Another way of asking this is: what **effect** does the wrong word have on the message? Is a message downright wrong in meaning, or simply wrong in style? For each route (cognates, transliterations, perceived equivalents, synonyms, associations and cores), there is a potential for each one to be dissonant in five different ways (although not every route and effect can be combined, and some route/effect combinations are more common than others).

Martin (1984) was the first to propose how (similar) words can be wrong in context, through four different types of lexical dissonance: semantic, stylistic, collocational and syntactic. To this, Hasselgren adds an additional category to cater to L2 speakers: invalid. Martin's account of effects is highly technical and does not explain how these effects work in an authentic context. Hasselgren's definitions of effects offer a more surmountable way of labeling dissonances.

5.2.2.1 Semantic

Hasselgren writes that: "**semantic dissonance** occurs when the meaning of the L2 item is completely different from the L1 source item" (Hasselgren 1993: 56, emphasis mine). This could be for instance semantic lexical transfer: when a student applies the senses of an L1 word to an L2 word. Since free production texts do not have an L1 source item, much like in the case of perceived equivalents, this definition is not helpful for my study. The only way to pinpoint semantic dissonance in free production texts is through context. So in my case, semantic dissonance is judged based on contexts where there is a mismatch between the word used, the possible senses that word could have and what I perceive to be the student's intended meaning based on context. In other words, in this study, semantic dissonance is a mismatch between L2 senses and L2 context, where the "right" word can or cannot be guessed through context. Examples of semantic dissonance found in my material include: *Her arms were* ***latent*** *on each side of her* (*lying*), or *it is not* ***immoral*** *to do some changes or improvements* (*bad/negative*).

5.2.2.2 Collocational

Collocational dissonance is defined as a word that is "inappropriate simply through disharmony with the words around it without being deviant in meaning or style" (Hasselgren 1994: 243). Unlike some semantic dissonances, collocational dissonances and their appropriate equivalent are discernible in free production texts because the dissonance is between the node word and its collocate (see section 3.2.3). Examples include: *An **evil** circle* (*vicious*) or to ***squeeze** a needle* (*inject*).

5.2.2.3 Stylistic/Connotational

Stylistic/connotational dissonance covers items that are semantically correct, but have an inappropriateness of style or connotation (Hasselgren 1994: 243). The reason why the categories stylistic and connotational are lumped together as one effect is that they frequently occur together. Here there is a mismatch between Hasselgren and Martins' taxonomies. Connotational words are labeled as semantic in nature by Martin. In this case I agree with Hasselgren – although there is a semantic aspect to connotation, the phrase *She **persisted** helping him* (continue) will strike a native speaker as having strange connotations more than being an actual misunderstanding of the sense of *persist.*

This study differentiates between informal, formal and connotational dissonance. **Informal example:** *Working with sick people may seem **gross*** (*unpleasant*). **Formal example:** *Important for a **human** to have a job* (*person*). **Connotational example:** *She was too busy [eating] the **cadaver*** (*dead body*). Words are considered connotational if the dissonance is perceived as being more connotational in nature than belonging to the formal/informal dichotomy.

5.2.2.4 Syntactic

Syntactic dissonance is dissonance where either the lexeme is right, but in the wrong word class, for instance: *they had an **argue*** (*argument*). Hasselgren chooses to omit syntactic cases where she considers dissonance to be grammatical rather than lexical in nature, such as adverbs lacking the -ly ending. I also attempt to weed out syntactic cases that are a reflection of grammatical rather than lexical features.

5.2.2.5 Invalid

This effect was created by Hasselgren to cater to items that do not exist in English. In her case, this means words that are created by students through transliteration. Hasselgren offers the example: *cold degrees* from Norwegian *kuldegrader* (Hasselgren 1994: 243). However, in my analysis, I also found non-words that were not the result of transliteration. Synonyms and associations also have non-words. Invalid seemed the most appropriate category for these non-words such as ***new-mown*** *grass* instead of *freshly cut,* or *the* ***predictment*** *of marrying is that you want to be together forever* instead of *prediction*.

5.2.3 Influence

A third question of interest in Hasselgren's study of lexical dissonance is: **why** do L2 speakers make mistakes? Or, what triggers a lexical dissonance (Hasselgren 1993: 63)? Here the term "influence" comes to play – "anything that is seen to cause a learner to make a wrong word choice" (Hasselgren 1993: 44). As mentioned in section 4.2, linguists commonly differentiate between two competing (and occasionally converging) influences: interlingual and intralingual influence. Hasselgren uses "divergence," which labels influence in terms of L1, L2, L1/L2 and "other" (for instance, cases of obvious avoidance). Since our data are so different, I found a need to create my own taxonomy for influences. This study divides influence in three ways: influence found in the L1-based routes (strong L1-influence, as described in 4.2.1.1), influence found in the L2-based (weak L1-influence as described in 4.2.1.2) and intralingual influence (see section 4.2.2).

Strong L1-influence is termed "strong" because it suggests that the dissonance is caused directly by a Norwegian L1 item. In the L1-based routes, there is always a possible L1 item that is the root of the dissonance. For instance, if a Norwegian student writes: *he is a very* ***high*** *man* instead of *tall*, it is reasonably certain that the dissonance is caused by the cognate *høy*. Strong L1-influence is the result of the word equivalence hypothesis: the L2 speaker has assumed an equivalence between an L1 and L2 item on some level, and the assumption that these aspects can be transferred from the L1 to the L2 has caused dissonance. Another cause is loan translation, or L2 items which are combined in ways only possible in the L1 (see 4.2.1.1 for more information)

Weak L1-influence is termed "weak" because it suggests an indirect possibility of the dissonance being caused by a Norwegian L1 item. Here it is not the influence of a particular word, but rather that the semantic space and general behavior of words is different between

the L1 and L2. These are normally instances of weak finer gridding (see section 4.2.1.2). For Norwegian speakers of English, the difference in stylistic Norwegian/English core word behavior is also considered weak L1-influence.

Intralingual (L2) influence covers instances where the dissonance cannot be explained by L1-influence, i.e. all dissonances that have no perceivable L1-influences are assumed to be L2-influenced. Examples include: *A smile crossed her* ***mouth*** (*lips*) or to ***quiz*** *a survey* (*take*). Neither of these instances can be explained by Norwegian, so we must conclude that there are other interferences at play, either within the L2 itself, or individual, personal mistakes that have caused the L2 speaker to encode or perceive the word wrongly.

5.2.4 Shortcomings with the taxonomy

The taxonomy poses several problems, particularly for free production texts. The first problem is determining L1-influence as a source of dissonance – there is less certainty in predicting L1 interference, since no L1 item has been provided. The second problem is overlapping found in some of the categories. Cognates/transliterations and synonyms /associations are labeled differently due to word form – for instance, a dissonance is a cognate if the L1 and L2 item are similar in form and meaning (*history* - *historie*) or a transliteration if the L1 and L2 item are similar in meaning but not form (*case* - *sak*). However, how does one judge when two words cease to be similar in form? For instance, *break* - *bryte* are cognates, but not very similar in form. So is *break* a cognate or a transliteration? In these cases it was personal judgment that decided which route was the most dominant/influential.

Another case in point is perceived equivalents. A dissonance is a perceived equivalent if the correct L1 item cannot be guessed in context (e.g. a **mortal** glance). However, a more skilled researcher may be able to understand what the L2 speaker meant in context, and thus label it as a synonymous dissonance. Lastly, some of the dissonances can have up to several routes and effects. For example, one student wrote: ***white-skinned*** *people*. This dissonance could be associational (*white* - *light*) if we view it as a mistake of word form. However, in Norwegian it is possible to say *hvithudet* (*white-skinned*), so this could be a transliterational dissonance. If the dissonance is considered transliterational, the correct L2 item is *white people*. Judging which dissonance belongs to which category in these ambiguous cases is up to the individual researcher. The more personal a judgment is, the less scientific it becomes, and studies conducted using the same taxonomy will have different results due to personal bias.

Additionally, Hasselgren's taxonomy can only reflect instances of coping with lexis when a dissonance arises. Routes and influence can also lead to successful lexical choices, such as if the L2 speaker chooses *high* for the collocation *high price* (*høy pris*). We do not have access to the student's mind, so there is no way of determining if the student has "guessed" the correct lexical item. One way to counter this is to count instances of positive lexical transfer (performance analysis) for a successful/wrong lexical choice ratio. Yet the same problem arises, since it is impossible to determine if the student has "guessed" their way to correct word choices as well. This is generally a problem when trying to etiologically explain errors.

6. METHODS AND MATERIAL

This thesis uses error analysis to examine distributional patterns of lexical dissonance. The taxonomy for classifying lexical dissonance outlined in the previous chapter will determine the distributional patterns based on how often dissonance occurs in these categories. The material in this study is taken from two separate corpora – one written by Norwegian students at an intermediate level of English, and one written by Norwegian students at an advanced level of English. The corpora are cross-sectional data, since they represent two groups in a specific period of time for the purpose of comparison. This section will discuss the methods and data used for this study, and additionally contrast some of the strengths and weaknesses that follow with the methods and data chosen.

6.1 Discussion of types of methods and data

6.1.1 Error Analysis

Error analysis is a tool used for investigating how learners acquire their L2 (Ellis 2008: 45). It is a method that has its origins in the field of SLA, and generally distinguishes between two main tasks: investigating gaps in knowledge, and investigating how speakers cope with these gaps (James 1998: 62). "Coping" in terms of lexis refers to communication strategies (ways of communicating when a learner lacks a word) and learning strategies (ways of remembering new words). Hasselgren believes that a third way of coping with errors is not by any strategic move, but is "more a matter of stumbling" when trying to pick the right word (Hasselgren 1993: 43). This is what we refer to as **routes**. Routes can lead to lexical dissonance. By identifying instances of "routes" that lead to lexical dissonance and examining how these routes are distributed in overall lexical errors, this study hopes to find indicators of how Norwegians cope with their English vocabulary.

Error analysis is undoubtably a useful tool, but has some limitations as well. Many have pointed out (e.g. Simensen (1998) and Johansson (2008)) that error analysis only *partially* reflects interlanguage. A full scope of what the learner both can and cannot do is a better representation of interlanguage. Such a method is called **performance analysis**, and studies the learner's overall performance, which would in this case include accurate use of lexis and investigation of pragmatic failures (Simensen 1998: 94). Other limitations of error

analysis include the bias of evaluation judgments (researchers will judge errors differently), lack of consensus of what an error is, and determining the source of an error (how can we know for sure if an error is caused by interlingual or intralingual influences?) (López, date of publication unknown).

6.1.2 Translation versus free production – pros and cons

There are two main types of data in error analysis: translation and free production. L2 translation data are texts which have been provided to the L2 speaker in their native tongue (in this case, Norwegian). The L2 speaker translates the text to their L2 (in this case, English), and the end result is a translation text. Translation texts are useful in investigation of learner language because the translation data reveals a) how the learner expresses L1 words in their L2, and consequently b) how the learner perceives the relationship between words in the L1 and L2. An advantage of translation data is that the L2 speaker's **intended meaning** (what they are supposed to express) is available to the researcher. If intended meaning is available, it is much easier to determine the L2 speaker's relationship between L1 and L2 words. For example, if an L2 speaker writes *discover* in a translation text, a quick peek at the original word *oppdage* will confirm that the L2 speaker has created an accurate 1:1 mapping between the L1 and L2 word. Such insights into intended meaning are of particular aid to error analysis, because sources of error can more easily be identified by seeing how the L2 speaker perceives the relationship between L1 and L2 words, i.e., what the L2 speaker perceives a certain L1 item to mean in their L2. Translation texts can therefore be said to more accurately identify L1-influenced errors. However, translation texts are not without fault. Firstly, being provided with L1 words is problematic on some level. The L2 speaker may have never used the word provided in either their L1 and L2, and being provided "cue" words means the L2 speaker is limited in selection of words. Furthermore, translation texts focus on the L1-L2 relationship (how an L2 word is interpreted in relation to an L1 word), and not how an L2 word is perceived in an L2-context.

In Hasselgren's 8 translation tasks, 15 examples of each were provided, but they are all a reflection of how 15 different students interpreted the number of words limited in eight tasks. This can be considered an advantage in that for those limited number of words, there are 15 different sets of data illustrating how learners cope with those particular words. However, it is also a disadvantage because being provided cue words does not reflect how an L2 speaker would naturally write about a theme, and the dissonances that follow free production without L1-items.

In the 1990s, L1 (Norwegian) was seen as a positive factor in the English language classroom (Simensen 1998: 123). Today, this is not the case. Translation tasks are therefore perhaps a discipline belonging to the past, when relationships between the L1 and L2 were more encouraged in the EFL classroom. With the rise of communicative competence and immersion techniques, it is becoming frequently more popular in Norway to rely less on L1-L2 glossaries and translation tasks, and instead allow students to use their own knowledge of English in the EFL classroom and build on their L2 vocabulary based on the words they already have acquired. Because of this, free production texts are more prudent to examine in Norway – Norwegian students are no longer trained in comparing their L1-L2 knowledge, and a comparison of L1-L2 knowledge is at any rate not the type of knowledge that is of interest to current EFL studies in Norway. However, free production texts also pose a number of problems in error analysis. First, there is no insight into the student's intended meaning, and the only way to deduce errors is through context. Errors are therefore more prone to interpretation based on the researcher. There is also a higher risk of avoidance of words the L2 speaker does not know, because there is no pressure to extract knowledge on specific L2 items. Lastly, there is less data on individual errors, because the participants can choose the words they use themselves, so it is much less likely that participants will have the same types of errors.

While Hasselgren (1993) uses translation data, this study draws on free production texts. To my knowledge, Hasselgren's study is the only study that attempts to find distributional patterns of lexical dissonances in Norwegian speakers of English. Initially, this study sought to replicate Hasselgren's, to compare and support or challenge her results. However, replicating Hasselgren's study was impractical due to the a) availability of translations texts, particularly for intermediate students, and b) how relevant and natural translation texts are for both intermediate and advanced Norwegian speakers of English, with regard to the focus of the Norwegian educational system.

6.1.4 Cross-sectional data

The two corpora used for this study are cross-sectional data, meaning a set of data at a specific point of time from a group that represents a whole. This is referred to as “snap-shot” corpora, because they are set at a certain time and do not show participants' progression over a longer period of time. The advantages of using such material is easy accessibility and comparisons of several subjects at a specific point in time. A disadvantage is that cross-sectional data cannot explain what will happen to the participants over time. Ellis calls this a

"static view of L2 acquisition" (Ellis 2008: 61). This study chooses to use cross-sectional data since longitudinal data for Norwegian speakers of English from intermediate to advanced is not available, and also because it wishes to draw comparisons between the two groups. Since this data is not longitudinal, it cannot be determined if there has been any "progression" from intermediate to advanced at an individual level. The participants are different, and the texts are written at different times, allowing for differences in data due to variables such as chance, different linguistic environment and individual differences.

6.2 The data and participants

6.2.1 The intermediate corpus

To my knowledge, there is no freely available corpus of intermediate Norwegian learners of English. This meant that I had to create my own corpus. Junior high school students seemed an ideal representative of intermediate Norwegian learners of English, so 18 winter (2012) mock exams in English were collected from a junior high school English class in the Oslo area. The students in this corpus are homogenous (Norwegian native tongue, none have lived abroad and all are fifteen years old). The texts are responses to national tests distributed to tenth graders across the country (see appendix 1 for tasks). The corpus consists of both argumentative and literary texts and are all free production. In preparation for the winter mock exam, students were given a booklet (called "FOCUS") with information regarding the various topics which would be covered on the exam, and could be used as a source. The mock exam lasted five hours. Students were allowed access to any aids, excluding communicative devices such as phones and the internet. The students also had access to spell check on their computer. Altogether the corpus comprises 24,732 words.

6.2.2 The advanced corpus

For Norwegian students of advanced English, the NICLE corpus was optimal. The NICLE corpus is a Norwegian subcorpus of ICLE (The International Corpus of Learner English). This corpus was initiated by Sylviane Granger of the Université Catholique de Louvain in 1990. The idea was to create comparable corpora of English texts by L2 learners of English with a specific L1 background. The texts for the corpora, produced by first-year university students, are mostly argumentative essays and about the same length, covering the same topics. ICLE contains approximately 4.5 million words and has sixteen subcorpora (McEnery & Hardie 2012: 82).

NICLE was compiled in 1992-2002 by Lynell Chvala and Stig Johansson at the University of Oslo. The total size of NICLE is around 300 essays and consists of approximately 212,000 words (Johanssson 2008: 117). Most essays are written at home with no time limit, with the possibility of using reference tools. The topics are mainly argumentative. For an overview of these topics, see appendix 2.

It is statistically unwise to compare a 200,000 word corpus with a corpus that barely spans 24,000 words, and also insurmountable to single-handedly analyze a corpus the size of NICLE. In order to compare NICLE to the intermediate corpus, the NICLE corpus had to be decreased in both size and participants, i.e., only a few select texts were chosen for the advanced corpus used in this study. For optimal comparison, a few filters had to be set when choosing which NICLE texts would be appropriate for the advanced corpus. Ideally, two comparable corpora should be the same size and have the same number of participants. This posed a problem, since the NICLE texts are much shorter. For this study, 39 texts (and thus participants) were chosen for the advanced corpus, amounting to a total of 24,299 words. The intermediate students had a certain background and access to tools. In order to match the conditions, and also ensure that the advanced students were truly advanced and as homogenous as possible, the texts from NICLE had to fulfill certain criteria. In the end, the two corpora had the following similarities and differences[8]:

Table 1. Comparable information about the two corpora used

	Intermediate Corpus	**Advanced Corpus**
L1	Norwegian	Norwegian
L2	English	English
Size of corpus	24,732 words	24, 299 words
No. of participants	18	39
Age of participants	14-15 years old	19-21 years old
Years of English at school	9.5 years	9-10 years
Genre	Argumentative / Literary	Argumentative
Type of texts	Mock exam	Take-home essay
Time limit	5 hours	No limit
Reference tools allowed	Yes	Yes

8 The intermediate and advanced students have nearly had the same number of years of English at school because of the new reform, which allowed the intermediate students to begin English from the first grade. However, this does not make the intermediate students less intermediate or the advanced students less advanced.

6.3 Procedure

6.3.1 Locating lexical errors

After assembling the two corpora, careful readings of each corpus were performed to identify lexical dissonance. Once a lexical dissonance was located, it was labeled for its route, effect and influence. All types of dissonance were recorded and can be found in appendix 3 (intermediate) and appendix 4 (advanced). Information on each individual type of dissonance can be found in the appendices, with the number of times each dissonance occurred, its route, effect, the right L2 item (if one could be discerned) and the possible L1 influential item if relevant.

The primary judger of lexical errors was me (a bilingual Norwegian/English speaker with L1 English). Each dissonance was only given one route/effect/influence. If a dissonance was for instance both transliterational and associational, the dissonance was awarded the route with the perceivably largest influence. In cases where this was ambiguous or difficult to judge, my supervisor (Signe Oksefjell Ebeling) aided me in the decision. Occasionally a dissonance fit none of the categories as described by Hasselgren (for instance my definition of perceived equivalents). In these cases, I altered the taxonomy to fit the data.

6.3.2 Some terms used in the data analysis

In the quantitative analysis, dissonances are represented in two ways: by tokens and types. **Tokens** account for the actual instances of word forms found. For instance, if there are 15 wrong instances of the word *rose*, this means that 15 instances of *rose* were found in total. A **type** is instances of unique words, i.e. if the word *rose* is wrongly chosen 15 times, it is still only one type of dissonance. In order to determine the relationship between tokens and types, a **repetition-ratio** is calculated. The repetition-ratio is calculated as follows:

$$\frac{\text{No. of tokens}}{\text{No. of types}} = \text{Repetition Ratio}$$

For example, if there are 15 tokens and 5 types found in cognates, the repetition-ratio is 3. Each type of cognate is repeated on average 3 times in the student texts. The higher the repetition ratio, the more repeated a mistake is. Repeated mistakes are viewed as more

problematic – e.g. if *rose* has a high repetition-ratio, this means *rose* is frequently repeated and problematic to L2 speakers.

6.3.3 Statistical Analysis

This study aims to find out whether the difference in number of intermediate and advanced dissonances is statistically significant. For example, 97 cognates are found in the intermediate texts and 46 cognates are found in the advanced texts. By using a statistical test to find p-values, it can be determined if the difference in number of cognates in these two groups is significant. A p-value of less than 0.05 indicates that the difference is significant.

Statistical significance is calculated by using a log-likelihood test. Paul Rayson's log-likelihood calculator[9] was used, since it is specifically meant for comparing corpus data. The number of dissonances found in each category are entered in the calculator (in this case, 97 intermediate and 46 advanced cognates) along with corpus size (24,732 words for the intermediate corpus and 24,299 words for the advanced corpus). The calculator then provides an LL-value (log-likelihood value). The p-value is determined based on the critical LL-value. They correspond in the following way[10]:

Table 2. LL-values to p-values

P-value	Critical LL-value
P < 0.05	3.84
P < 0.01	6.63
P < 0.001	10.83
P < 0.0001	15.13

9 http://ucrel.lancs.ac.uk/llwizard.html accessed 01.05.2013

10 Chart taken from: http://ucrel.lancs.ac.uk/llwizard.html , accessed 01.05.2013

7. DATA ANALYSIS

This section will present the findings of an error analysis of two corpora (intermediate and advanced) based on the taxonomy of lexical dissonance. First a general overview will be given, with a summary of some of the total statistics for in this study. Second, the two sets of data (intermediate and advanced) will be contrasted for routes, effects, the route/effect combination and influence, with discussions of some of the results.

7.1 General Overview

Table 3. Number of total dissonances found in the Norwegian learner data (intermediate and advanced)

	No. of tokens	No. of types	No. of L1 influences
Dissonances	1077	476	603

From a body of text consisting of 49,031 words, 1077 of the lexical words are dissonant in some way. Two percent of the entire corpus, or 1 in every 50 words is dissonant (although not all words in a text are lexical). These 1077 dissonances are represented by 476 different types. The repetition-ratio is 2.3, meaning each dissonance is on average repeated a little more than twice. Fifty-six percent of the tokens have a perceivable L1-influence.

The routes and effects are distributed as follows:

Figure 5: Distribution of routes in corpora

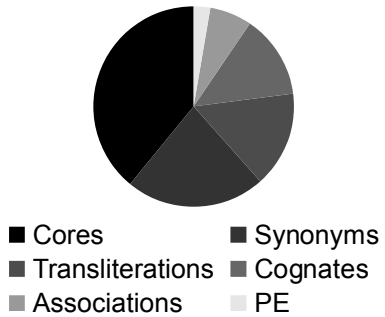
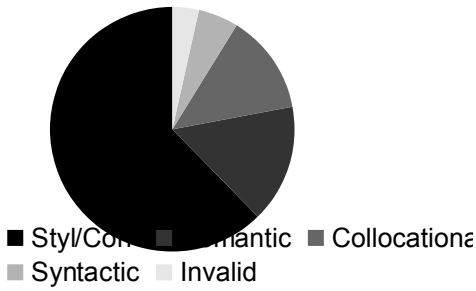


Figure 6: Distribution of effects in corpora



The routes and effects have the following percentage distribution in descending order:

Figure 7: Percentage distribution of routes and effects in corpora

Routes

Cores (39%) > Synonyms (22%) Transliterations (16%) > Cognates (13%) > Associations (7%) > PEs (3%)

Effects

Styl/Con (62%) > Semantic (16%) > Collocational (13%) > Syntactic (5%) > Invalid (4%)

7.1.1 Routes

Cores is the largest route, indicating that the largest lexical coping mechanism for Norwegian learners of English is using a non-specific vocabulary (e.g. *like* instead of *enjoy* or *appreciate*). Cores represent as much as 4 in 10 wrong words.

Synonyms is the second largest route, showing that dissonance in Norwegian learners of English will often be caused by picking a somewhat related, but wrong word (*seemingly* for *apparently* or *going* for *walking*). Every 1 in 5 wrong words is caused by this.

Transliterations is the third largest route, giving evidence of literal translations for a little more than 1 in every 6 wrong words (***get away*** (*få vekk*) instead of *remove*, to ***make*** *a party* (*lage fest*) instead of *throw a party*)

Cognates is the fourth largest route, illustrating that interference of similar words between two languages causes a wrong word choice for 1 in every 6 dissonances. Examples include *to* ***mean*** *that* (*å mene at*) instead of to *believe/be of the opinion*, or to ***have*** *the impression* (*ha inntrykket*) instead of *be under the impression*.

Associations is the second smallest route, indicating that every 1 in 14 dissonances is caused by a confusion of word form (***defiantly*** for *definitely*, or *the ocean* ***raised*** for the *the ocean rose*).

Perceived equivalents is the smallest route, accounting for 1 in every 33 dissonances. This would suggest that for every 33rd dissonant word, the L2 speaker did not know the word form or meaning of what s/he intended to say at all, or had completely misunderstood the meaning

of the word (*To give a **withering** look* (unknown intended meaning), or *to **evolve** something* instead of *develop*).

7.1.2 Effects

Stylistic/Connotational is the largest effect, suggesting that the most common lack of knowledge for wrong words chosen is related to register. Stylistic/Connotational dissonance is evident in over half of all dissonances, and includes stylistically inappropriate (most often informal) words such as ***crazy*** for *mentally ill*, or ***get*** for *receive/find/obtain*.

Semantic is the second largest effect, implying that for approximately 1 in every 6 dissonances, L2 speakers are not fully aware of the senses of the word they are using. Examples include: *The **divisions** of a hospital* instead of *wards*, or *a **wood*** for *a forest*.

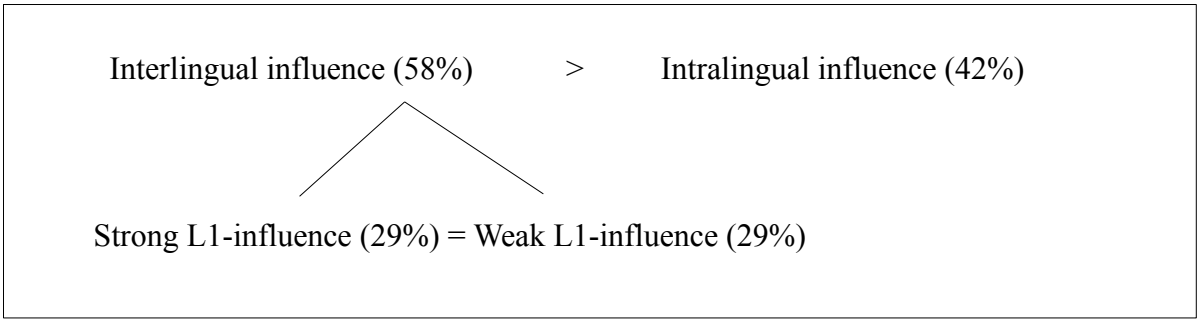
Collocation is the third largest effect, illustrating that lack of word combination knowledge can explain 1 in every 6 dissonances. Examples are: ***Choose** the right decision* instead of *make,* or ***increase** your look* instead of *enhance*.

Syntactic is the second smallest effect, showing that confusion between words in word families or case relations occurs in 1 in every 20 wrong words. Examples: ***Get** born* instead of *be born* or a *near-**dying** accident* instead of *near-death*.

Invalid is the smallest effect, suggesting that in slightly less than 1 in every 20 wrong words, Norwegian speakers of English will either translate a Norwegian concept that does not exist in English (*life-lie* from Norwegian *livsløgn*, which means "illusions that a person may build their life around") or somehow create a word that does not exist (***consum*** for *consumption*).

7.1.3 Influences

Figure 8: Distribution of influence in corpora



Interlingual influence is the largest type of influence (58%), and can be subdivided into strong L1-influence and weak L1-influence.

Intralingual influence is the smallest type of influence (42%), suggesting that in slightly less than half of all dissonances, there is no perceivable trace of influence from the L1.

7.2 Intermediate and Advanced Comparisons

This study seeks not only to find general distributions of lexical dissonance in Norwegian speakers of English, but also to compare intermediate and advanced students for similarities/differences. This section will examine some of these similarities and differences in routes, effect, the route/effect combination and influences. First a general overview of each corpus and its dissonances is presented:

Table 4. Number of total dissonances found in intermediate versus advanced corpora

	No. of tokens	No. of types	No. of L1 influences
Intermediate Students	704	301	403
Advanced Students	373	212	200

Table 5. Comparable information of dissonances in the intermediate and advanced corpus

	Intermediate	Advanced
Size of corpus	24,732 words	24,299 words
No. of participants	18	39
Percentage of dissonance*	2.8%	1.5%
Repetition-ratio**	2.3	1.8
Percentage of L1-influence	57%	54%

* The percentage of dissonances also includes grammatical words, so in reality, the percentage of dissonances is much higher

** The repetition-ratio number indicates how many times a dissonance is repeated on average

For the intermediate students, approximately 1 in every 35 words is dissonant, while for the advanced students, this is 1 in every 67 words. The repetition-ratio is higher in the intermediate students, indicating that they are more prone to repeating the same type of mistakes. There is also more perceivable L1-influence in the intermediate corpus.

There is a clear difference between the number of dissonances found in the intermediate corpus versus the advanced corpus. To see if these two sets of dissonance (704 intermediate tokens and 373 advanced tokens) were statistically significant, a log-likelihood test was conducted. The log-likelihood for difference in number of dissonances between the corpora gave an **LL-value of 97.63**, meaning the difference in number of dissonances was statistically significant to the 99.99^{th} percentile. This gives a p-value of $\mathbf{p < 0.0001}$. The intermediate group can therefore be said to have significantly more dissonances than the advanced group.

7.2.1 Routes

This section will show how the dissonances are distributed across the six routes. This distribution can be divided into tokens (how many individual dissonances were found per route) and types (how many different types of dissonances were found per route). There are therefore two ways of determining which route is largest: through counting tokens (which route has the most amount of dissonances) and types (which route has the most amount of different types of dissonances). Additionally, the repetition-ratio can suggest which routes have the most repeated types of dissonance. In order to determine if the distribution of routes

between intermediate and advanced is statistically significant, a log-likelihood test will be conducted. Lastly, there will be a comparison of the distribution of routes within the corpora themselves. A comparison of distribution can illustrate if the proportion of distribution is different or the same. For instance, the intermediate students have 18 more synonymous dissonances than the advanced students. However if we look at the proportion of routes within the intermediate and advanced corpora, synonyms account for 19% of intermediate routes and 30% of advanced routes. This means that although the intermediate students have more instances of synonyms, they rely less on this route than the advanced students proportionally speaking.

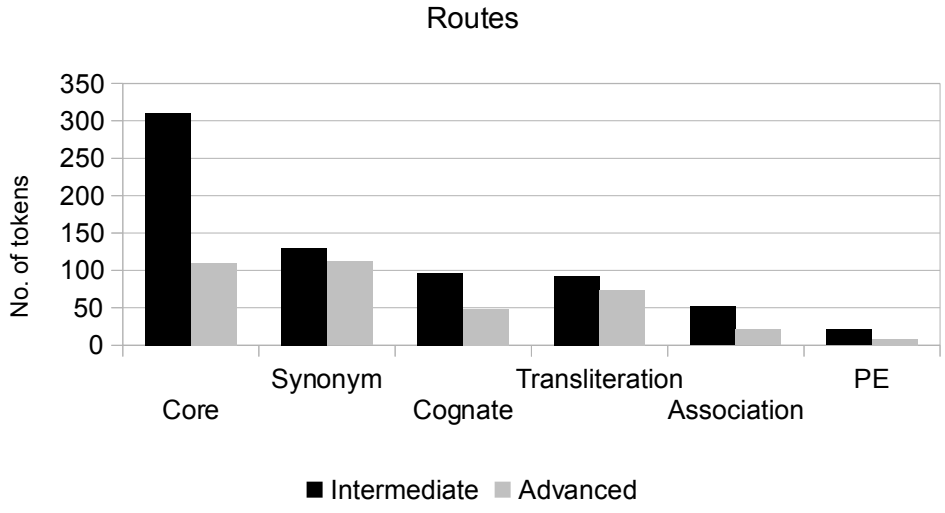
Here is the total distribution of routes for intermediate and advanced students.

Table 6. Total number of tokens and types in routes for intermediate and advanced corpora

Route	Intermediate Students		Advanced Students	
	No. of tokens	No. of types	No. of tokens	No. of types
Cognate	97	45	48	25
Transliteration	93	68	74	53
Perceived Equivalent	21	14	8	8
Synonym	130	102	112	91
Association	52	47	21	19
Core	311	25	110	16
Total	**704**	**301**	**373**	**212**

7.2.1.1 Comparison of tokens

Figure 9: Ranking of routes in tokens for intermediate and advanced students

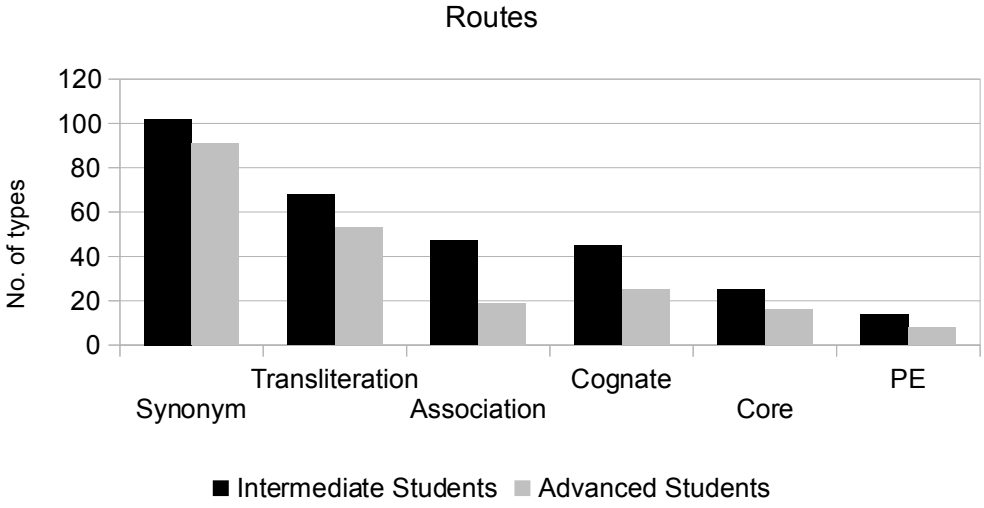


The token ranking shows that cores constitute by far the most tokens for intermediate students. Cores are high for advanced students, but are surpassed slightly by synonyms (by two tokens). The fact that students choose the wrong word (synonyms) just as much as they inappropriately use basic words suggests that advanced students will more often pick the wrong word instead of "clinging" to core words. Intermediate students, however, prefer the safety of core words. If we assume that advanced students have a larger vocabulary, a source of dissonance could perhaps be the choices that come with coping with a larger vocabulary. The result will be dissonances from the many choices an advanced L2 speaker faces when having to choose the "right" word in context.

For both groups synonyms rank high (18% for intermediate and 30% for advanced), and association and perceived equivalents rank low. In the advanced student texts, transliterations are preferred to cognates. The preference for cognates and transliterations appears to be equal for intermediate students.

7.2.1.2 Comparison of types

Figure 10: Ranking of routes in tokens for intermediate and advanced students

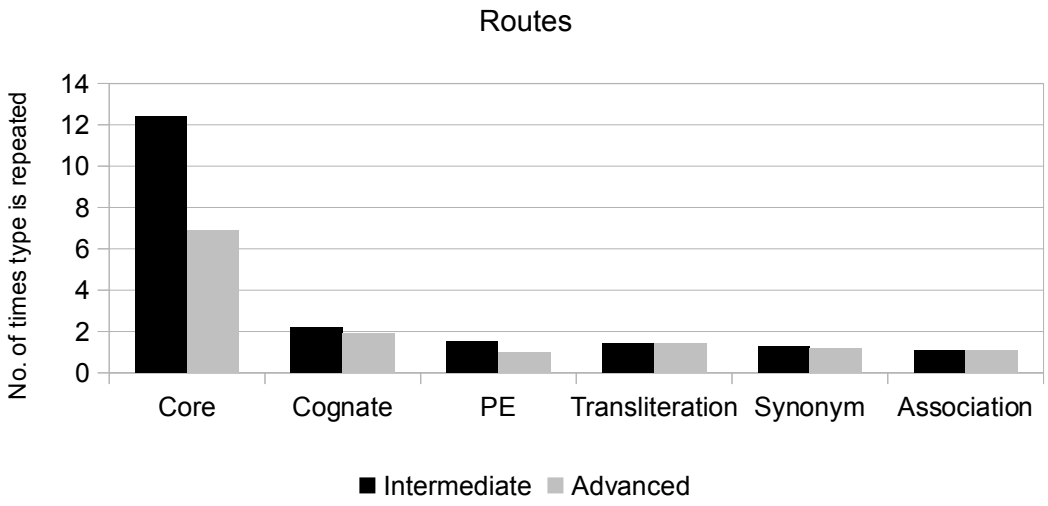


A type ranking illustrates that although cores have the most tokens, the most varied type of dissonance is synonyms. Cores may have the most dissonances in total, but they are often the same types of dissonance that are repeated several times, thus making cores fall from the largest token category to the second smallest type category. Synonyms and transliterations have many more different types; therefore generalizing and categorizing why students make synonymous and transliterational dissonance is much harder, since the types of dissonances students make in these categories are so varied. While cognates and transliterations have

roughly the same amount of tokens (particularly in the intermediate group), transliteration has many more types in both groups. This can be caused by several factors: a) There are fewer cognates than transliterations, b) only certain types of cognates are confusing to Norwegian speakers of English, while transliteration is generally confusing, c) the English cognates that confuse Norwegian speakers of English are frequent in the L1 and/or d) transliteration is viewed as more acceptable and transferable than cognates.

7.2.1.3 The repetition-ratio

Figure 11: Repetition-ratio for routes in intermediate and advanced students



The repetition-ratio shows that cores are indisputably the highest repeated dissonances. Core dissonances are repeated between 7 (advanced) to 12 (intermediate) times on average. In tokens it was suggested that several core items are "replaced" by synonymous dissonance in the advanced students. An overview of the repetition-ratio reinforces that not only do advanced students have fewer core dissonances, but the core dissonances found in advanced texts are also repeated considerably less.

Cognates are the second most repeated route, with each dissonance repeated on average two times for both groups. This would suggest that the same types of cognates are confusing to Norwegian speakers of English (e.g. *mean - mene*). Due to the high repetition-ratio for cores and cognates, lists of the most common cores and cognates can therefore be made for Norwegian speakers of English, so they can be aware of their most frequent dissonances.

Perceived equivalents, transliterations, synonyms and associations all have only around one dissonance each per type, suggesting that students who use these routes do not tend to repeat the same dissonance. The consequence of this is that creating lists for these types of dissonance for these four routes is much more challenging, since many of these dissonances are idiosyncratic (pertaining to the individual L2 speaker).

7.2.1.4 Statistical significance

There are clear differences in the number of dissonances found in routes for intermediate and advanced students. In order to find out if the difference between these two groups was statistically significant, a log-likelihood test was conducted on each route:

Table 7. LL-values and p-values for statistical differences between intermediate and advanced routes

Route	LL-value	P-value	Result
Cores	96.47	P < 0.0001	√
Cognates	16.03	P < 0.0001	√
Associations	13.05	P < 0.001	√
PEs	5.81	P < 0.05	√
Transliterations	1.84	P > 0.05	x
Synonyms	1.04	P > 0.05	x

√ = difference between intermediate and advanced is statistically significant

x = difference between intermediate and advanced is not statistically significant

The results of the log-likelihood test suggest that there is a statistically significant difference in cores, cognates, associations and perceived equivalents between the intermediate and advanced student texts. There are far more dissonances for the intermediate students for these routes. Particularly “cores” gives a high LL-value, showing that this route is by far the largest divergence between the two groups. However, the differences in transliterations and synonyms in the two groups are not statistically significant, indicating that advanced students have enough dissonances in these two routes for there not to be a significant difference between the two groups.

7.2.1.5 Comparative distributions

Figure 12: Comparative distribution of routes in intermediate students

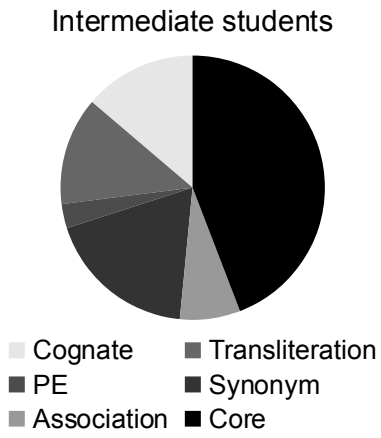


Figure 13: Comparative distribution of routes in intermediate students

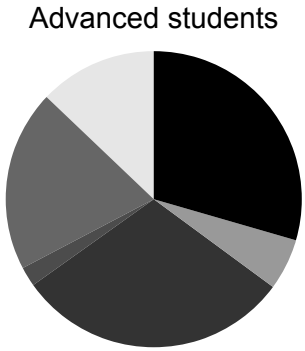


Figure 14: Comparative distribution of routes for intermediate and advanced students

Intermediate

Cores (44%) > Synonyms (19%) > Cognates (14%) > Transliteration (13%) > Association (7%) > PEs (3%)

Advanced

Synonyms (30%) > Cores (29%) > Transliteration (20%) > Cognates (13%) > Association (6%) > PEs (2%)

The comparative distribution confirms that, proportionally speaking, intermediate and advanced students have different ways of choosing the wrong word. The difference between distribution of cores within the two groups is 15%, confirmed by the log-likelihood test to be the most statistically different way of choosing the wrong word. Proportionally speaking, the advanced students also show a preference for synonyms and transliteration compared to the intermediate.

7.2.1.6 Summary of routes

Cores are the preferred route for intermediate students, while this preference is distributed among cores, synonyms and transliterations among the advanced students. This is reflected both by the log-likelihood test, which showed cores to be the highest LL-value and synonyms and transliterations to not be significantly different, and the comparative distribution, which has different percentage distributions for the three routes. Not only are there fewer core tokens for the advanced students, but core dissonances are also less repeated by advanced students, as illustrated by the low ranking in the repetition-ratio. The growing preference for synonyms in advanced students may be explained by the fact that a larger vocabulary can be confusing to an advanced student, and the trouble of the L2 lies less in clinging to "known" words and more in coping with the copious amounts of choices an advanced student faces when picking the "right" word. The differences between intermediate and advanced routes suggest that the more vocabulary a student has, the less s/he will rely on cores. However, this larger vocabulary that comes with being at an advanced level is still confusing, as it constitutes more choices for the L2 speaker. Synonymous dissonance generally has more types than cores, indicating that core dissonances can be more easily "fixed" in the EFL classroom by examining core word mistakes that are often repeated.

When it comes to cognates and transliterations, they are preferred equally by the intermediate students, while transliterations are preferred by advanced speakers. This suggests that advanced students prefer to "translate" Norwegian phrases, word combinations and stylistic elements, but why this is a preferred route among more advanced speakers is not clear. Cognate dissonances have a higher repetition-ratio, meaning they are generally repeated more often than transliterational ones. This suggests that a) certain cognates are more frequently misunderstood by Norwegian speakers of English, b) the cognates which are misunderstood are vital to the structure of the L2 and/or c) transliterations are seen as more acceptable.

The routes "association" and "perceived equivalents" are the smallest token groups for both intermediate and advanced students. However, "association" has a high ranking in types. The reason for this is that these mistakes are highly individual, so the same dissonance is almost never repeated. The same can be said of perceived equivalents. These two routes are both small and difficult to categorize, and proved to be the least problematic for students at both an intermediate and advanced level.

The log-likelihood test confirms a statistically significant difference between all routes

except transliterations and synonyms. This indicates that the other routes are used much more by the intermediate students, while transliterations and synonyms are popular among the advanced students as well. The comparative distribution shows that routes have different proportional distributions within the two groups, i.e. not only are there many more dissonances in the intermediate texts, but the way they are distributed across routes is proportionally different.

7.2.2 Effects

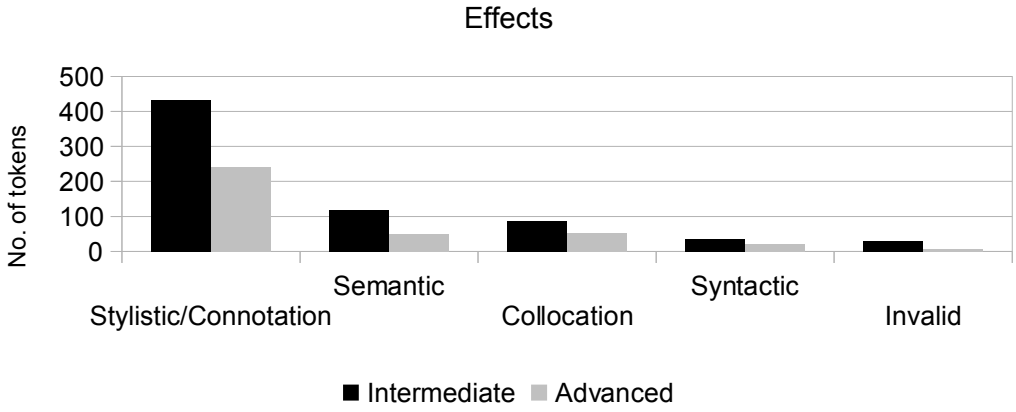
This section will show how the dissonances are distributed across the six effects. This distribution can also be divided into tokens (how many individual dissonances were found per effect) and types (how many different types of dissonances were found per effect). Effects will be treated the same as routes: first the token rank will be presented, then the type rank, repetition ratio, log-likelihood test for statistical significance and lastly the comparative distribution.

Table 8. Total number of tokens and types in effects for intermediate versus advanced students

Effect	Intermediate Students		Advanced Students	
	No. of tokens	No. of types	No. of tokens	No. of types
Semantic	117	72	51	40
Collocation	88	69	54	40
Style/Connotation	432	101	240	106
Syntactic	36	30	21	19
Invalid	31	29	7	7
Total	**704**	**301**	**373**	**212**

7.2.2.1 Comparison of tokens

Figure 15: Ranking of effects in tokens for intermediate and advanced students

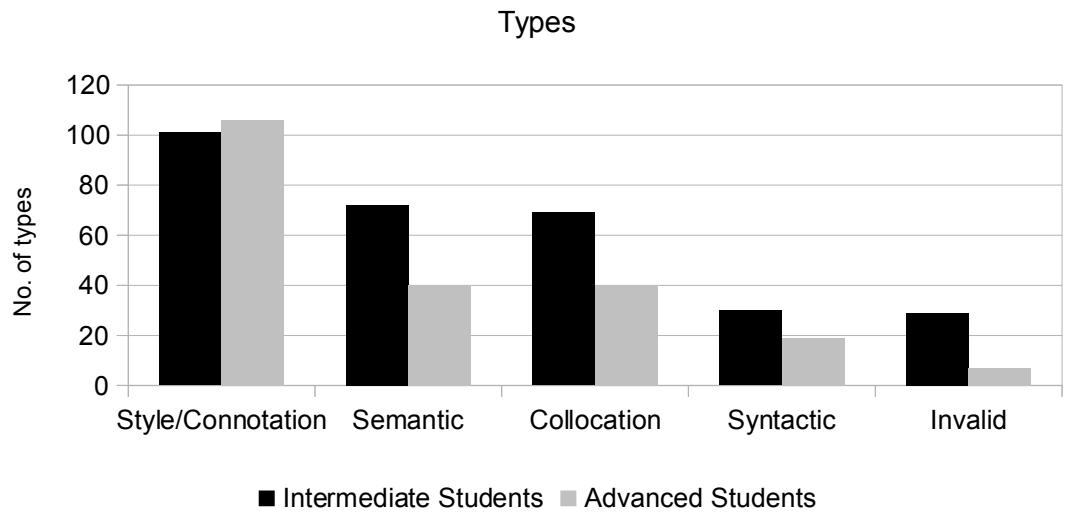


Stylistic/connotational dissonance constitutes by far the largest category in both intermediate and advanced students. The distribution of stylistic versus semantic effects suggests that Norwegian students of English struggle more with how words are stylistically marked than their actual meanings. This does not mean that semantic dissonances are non-existent – they are somewhere between 14% (advanced) and 16% (intermediate) of all dissonances. In addition, as mentioned in section 6.1.2, a weakness with free production texts is that there is no access to the speaker's intended meaning – i.e., if a particular word works in a certain context, there is no way of knowing if the student was thinking of another item or concept (for example, if a students writes *she broke a **cup*** but is thinking of a vase). So it must be noted that this effect could be much higher, and would probably be much more evident if translation texts had been examined, as confirmed by Hasselgren, who found semantic to be the largest effect.

The third largest effect is collocational dissonance for intermediate students, while it is tied with semantic dissonance for advanced students, indicating that advanced students struggle equally with meaning and the combination of different words. Syntactic and invalid are small in both groups – syntactic dissonance is between 5% (intermediate) and 6% (advanced), while invalid is between 2% (advanced) and 4% (intermediate).

7.2.2.2 Comparison of types

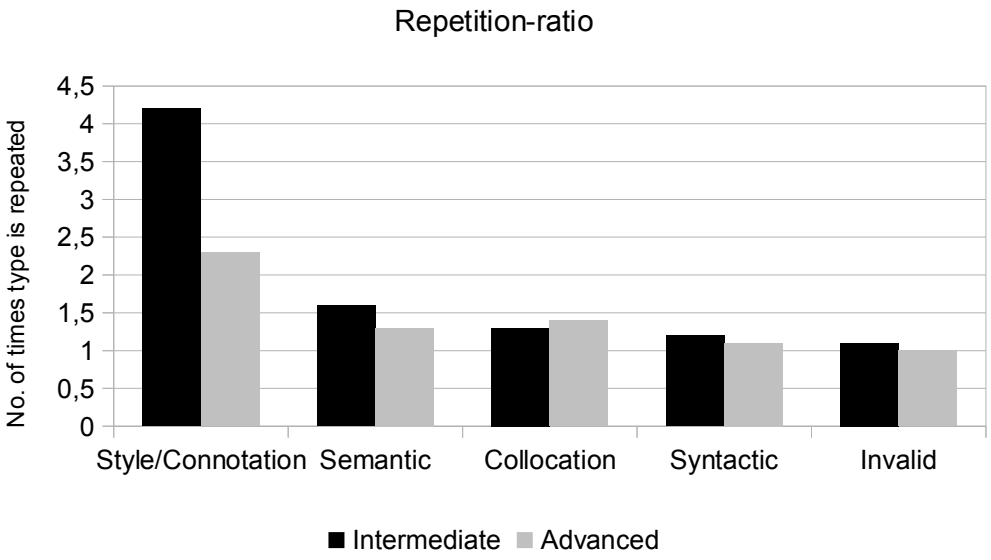
Figure 16: Ranking of effects in types for intermediate and advanced students



Types inform us that the categories hold the same rank in both tokens and types. However, particularly stylistic/connotational dissonances have many more different types in the advanced students. Collocation also shows a smaller gap between intermediate and advanced students in types, suggesting that stylistic/connotational and collocational dissonances in advanced students are more varied – there are for instance over 100 different types of stylistic dissonances in both groups.

7.2.2.3 The repetition-ratio

Figure 17: Repetition-ratio of effects for intermediate and advanced students



The repetition-ratio suggests that stylistic/connotational dissonances are the highest repeated dissonances in both intermediate and advanced students. A reason for this could be the inclusion of lexical teddy bears – words that are repeated several times and are stylistically inappropriate are labeled as "cores," even though they may also be transliterational or synonymous in nature (see section 5.2.1.6.4). Semantic and collocational dissonances also re-occur, as for instance each semantic dissonance occurs on average 1.5 times for intermediate students.

7.2.2.4 Statistical significance

The log-likelihood test showed which of the effects were statistically significant.

Table 9. LL-values and p-values for statistical differences between intermediate and advanced effects

Effect	LL-value	P-value	Result
Styl/Con	53.26	P < 0.0001	√
Semantic	25.49	P < 0.0001	√
Invalid	21.91	P < 0.0001	√
Collocation	7.63	P < 0.01	√
Syntactic	3.73	P > 0.05	x

√ = difference between intermediate and advanced is statistically significant

x = difference between intermediate and advanced is not statistically significant

The results of the log-likelihood test suggest that there is a statistically significant difference in all effects but syntactic. Moreover, advanced students are shown to struggle much less with stylistic and semantic dissonance than the intermediate students, whereas the gap between advanced/intermediate dissonance in the collocational and syntactic categories is much smaller (i.e. both groups struggle with these two effects). However, while collocational dissonance is still found to produce a statistically significant result, this is not the case for syntactic dissonance.

7.2.2.5 Comparative distribution

Figure 18: Comparative distribution of effects for intermediate students

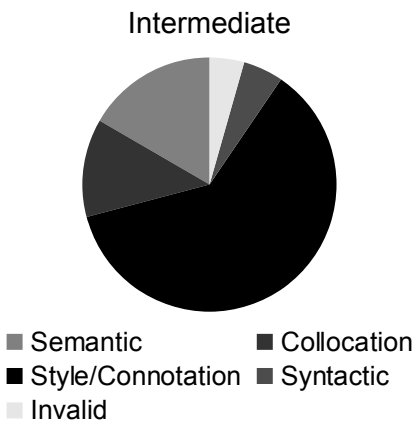


Figure 19: Comparative distribution of effects for advanced students

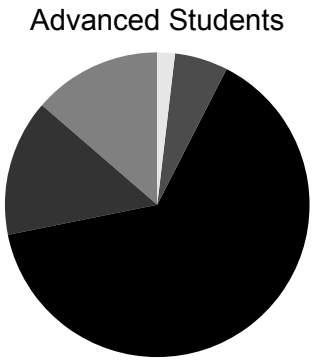
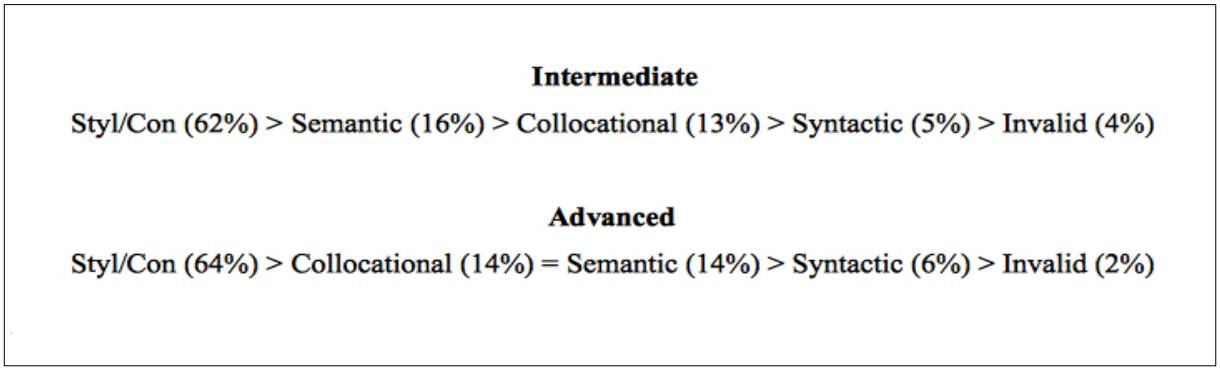


Figure 20: Comparative distribution of effects for intermediate advanced students



The comparative distribution gives some interesting insight: it appears that although each effect is larger in intermediate students, the distribution within each corpus is almost identical in the two groups, i.e. the advanced distribution of effects is only proportionally smaller. This would suggest that although intermediate and advanced students have a different distribution of routes, or ways of choosing the wrong word (see figures 12 and 13), the incorrect hypotheses they create about these wrong words (or the lack of knowledge they have of the wrong word) is the same. This may suggest, in concordance with Schmitt's word knowledge continuum (see figures 1, 2 and 3), that some aspects of word knowledge are more difficult than others, e.g. that stylistic knowledge is hard to acquire regardless of level of English.

7.2.2.6 Summary of effects

Stylistic/connotational dissonance is unequivocally the largest effect in both groups, constituting well over half of all dissonances. This suggests that above all, Norwegian students of English are unaware of what types of registers English words belong to. The repetition-ratio illustrates that stylistic dissonances are also the ones that are most repeated.

Semantic/collocational dissonance also proves to be problematic for Norwegian speakers of English, both in terms of types and tokens. Unlike the syntactic and invalid categories, semantic, collocational and stylistic dissonances are repeated up to several times, and these dissonances are reflected in several student texts. Syntactic and invalid dissonances are more idiosyncratic (they depend on how the individual L2 speaker perceives their L2 and cannot be easily categorized) and are the smallest effects for both intermediate and advanced students.

The log-likelihood test gives a significant difference in all effects but syntactic. The comparative distribution illustrates that although each group (besides syntactic) is significantly larger for the intermediate students, the proportion of effects is distributed equally in both groups. This suggests that when L2 students use words they only have partial knowledge of, the way in which these words are wrong are the same for both the intermediate and advanced learners.

7.2.3 The route/effect relationship

All routes must have an effect, or a reason **why** they are dissonant. Closer inspection of routes and effects will inevitably lead to the conclusion that there is a relationship between the two: some effects will combine more often with certain routes. This relationship is of interest because it can give insight into how effects are distributed in routes, i.e., what types of incorrect hypotheses are formed for which wrongly chosen word? For example, are cognates mostly wrong because there is semantic, collocational or stylistic dissonance?

Table 10. The route/effect combination of all dissonances for intermediate students

Intermediate Students						
Route	Semantic	Collocational	Styl/Con	Syntactic	Invalid	Total
Cognates	42	22	17	5	11	97
Transliteration	15	22	43	3	10	93
PE	19				2	21
Synonyms	24	35	62	7	2	130
Associations	16	9		21	6	52
Cores	1		310			311
Total	117	88	432	36	31	704

Table 11. The route/effect combination of all dissonances for advanced students

Advanced Students						
Route	Semantic	Collocational	Styl/Con	Syntactic	Invalid	Total
Cognates	12	12	18	5	1	48
Transliteration	8	15	45	3	3	74
PE	8					8
Synonyms	13	25	68	5	1	112
Associations	9	2		8	2	21
Cores	1		109			110
Total	51	54	240	21	7	373

= no existing route/effect combination for the particular corpus

7.2.3.1 Distribution of effects in routes

By examining how effects are distributed in each individual route, it is possible to see which effects are proportionally larger in each route. This section will illustrate the distribution of effects in routes through stacked bar charts, in addition to providing possible explanations for why some effects are more reoccurring than others. The routes "cores" and "synonyms" are much larger than the other routes, and therefore difficult to make the comparison in one single chart, so the distribution will be divided into the following: cognates, transliteration, perceived equivalents and associations in one chart, and cores and synonyms in another.

Figure 21: Distribution of effects in routes for intermediate students

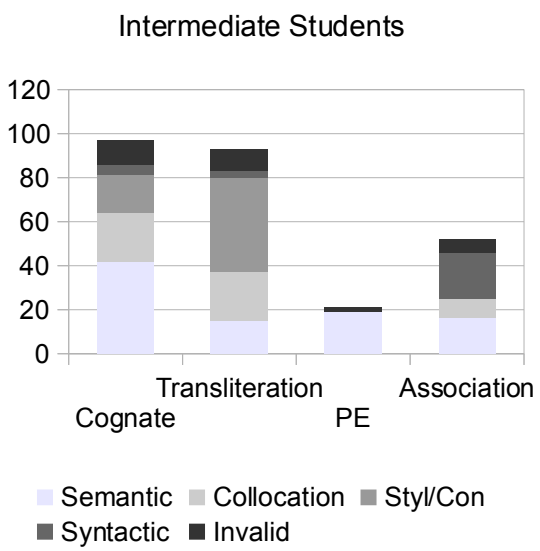
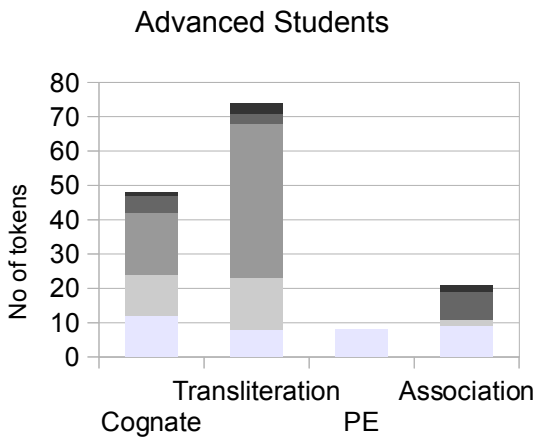


Figure 22: Distribution of effects in routes for advanced students



All effects are present in each route, apart from perceived equivalents (and style/connotation in association). Semantic dissonance is distributed approximately equally in all four routes apart from intermediate cognates, which show a higher degree of semantic dissonance. This then indicates that intermediate students are more prone to using cognates when they do not know the sense of the English word. This is a case for the **semantic equivalence hypothesis**, i.e. that the learner assumes the L1 and the L2 item are the same, without actually knowing how to use the L2 word (see section 4.2.1.1).

Collocational dissonance is discernible only in cognates, transliterations and associations. This is due to the semantic nature of perceived equivalents – if the student does not know the word they are writing, how can they possibly combine it with other words?

Cognates and transliterations are common routes for tackling lack of collocational knowledge for intermediate students, that is, it conforms to the **word equivalence hypothesis**, where learners will rely on their knowledge of L1 word combinations and apply this to L2 words (see section 4.2.1.1).

Stylistic dissonance is found only in the L1-based routes (cognates and transliterations). It would be rare to find two L2 items that are similar to each other both in word form and meaning while being stylistically different, so therefore no stylistically associational dissonances could be found. Cognates have less stylistic dissonance than transliterations. This could also be due to the fact that it is rarer to find an L1 and L2 item that are similar in word form and meaning, yet differ stylistically than an L1 and L2 item that are **not** similar in word form while being similar in meaning and differ stylistically. This could be explained by the fact that cognates are a narrower route (i.e. that there are fewer cognates than possible transliterations. See section 8.1 for more info). However, this could also be because students at both an intermediate and advanced level may simply prefer transferring stylistic knowledge between words that are similar in meaning and **not** form.

Syntactic dissonance occurs in cognates, transliterations and associations, and is highest for associations. This is because word family dissonances are taken into account, i.e. if a sentence has the right base form of a word, but it belongs to the wrong word class (*love is **difference*** instead of *different*). Such dissonances are not L1-based and therefore labeled as associational. These sorts of associational dissonances are the largest source of syntactic dissonance. It is also possible for Norwegian speakers of English to be syntactically confused by cognates and transliteration. Example: *to describe [something] very **good*** (*forklare [noe] veldig godt*). Such instances are not prominent in cognates (5 tokens for both intermediate and advanced students) or transliterations (3 tokens for both intermediate and advanced students), but do exist.

Invalid dissonance is found in cognates, transliterations and associations, and is trademark of the intermediate students. This could be because a) intermediate students may not be aware that several concepts only exist in Norwegian, such as *å gi et vink* (to give a wink) are not transferable to English and b) intermediate students are more likely to create nonsensical words (e.g. *burning damage*).

Figure 23: Distribution of effects in synonyms and cores for intermediate students

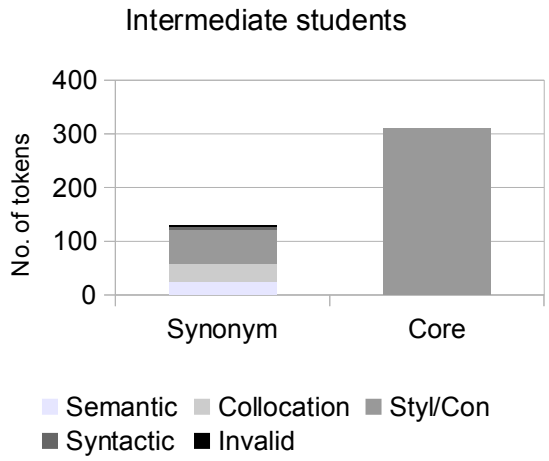
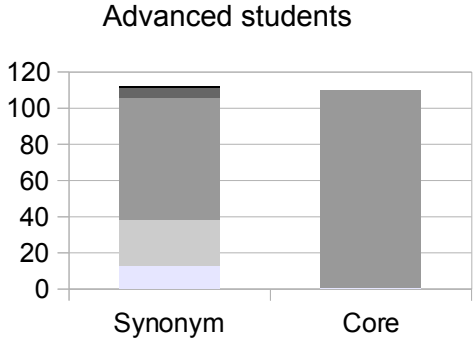


Figure 24: Distribution of effects in synonyms and cores for advanced students



Cores appear to be almost exclusively stylistically dissonant, with the exception of 1 semantic token per corpus. This is due to the nature of cores – students will seldom misunderstand the meaning of *good* or *nice*, and grouping these words with other words will rarely lead to collocational dissonance (*a* ***good/nice*** *time*). The reason why such words are dissonant is because they are not appropriate in a formal register. The result is that they are stylistically dissonant. Synonyms, on the other hand, can be misunderstood on several levels. Style/connotation is also the largest effect in this route, and the reason for this is the use of stylistically informal words which are **not** core words because they are not neutral and normally have specific connotations, such as ***crazy*** *person* (*mentally ill*) or ***old people*** (*senior citizens*). Such stylistic dissonance accounts for between 48% (intermediate) to 61% (advanced) of all synonyms. Since these stylistic dissonances are not from an L1-based route, "synonyms" suggests that Norwegian speakers of English use stylistically informal words they have acquired outside of their L1, i.e. that they are exposed to informal English, and/or feel more comfortable using a particular word. Synonyms can also be misunderstood semantically (***essential*** *cause* (*main cause*)) and collocationally. This indicates that Norwegian speakers of English will not always rely on their L1 for meaning and to create word combinations.

Synonyms and cores are L2-based routes since they rarely have direct influence from the L1. This means that the effects in cores and synonyms reflect the difficulties of coping with an L2 vocabulary. Some of these difficulties could be the choices that come with having to pick words from an incomplete mental lexicon, confusing synonyms that are often linked

semantically (*essential - main*) and the reluctance to use more specific words that would be more appropriate in context instead of basic words, as in the case of cores and lexical teddy bears (for instances overusing *boring* instead of using more specific and marked words such as *dull/ mundane/ meaningless*).

7.2.3.2 Summary of effects in routes

All routes appear to have semantic misunderstandings, apart from cores, since this route rarely includes words that can be misunderstood. Lack of semantic knowledge is remedied mainly by cognates in the intermediate students, and also transliterations and synonyms to some extent. The distribution of semantic effects in routes is more even among the advanced students, suggesting there is no distinct preference for L1-based routes when coping with gaps in semantic knowledge at an advanced level.

Collocational dissonance is found in cognates, transliterations, associations and synonyms. Synonyms have the largest amount of collocational dissonance (particularly in the advanced students), indicating that students will often choose the wrong synonym, with no clear preference or pattern. However, if we combine cognate and transliterational collocational dissonance, they will account for exactly 50% of collocations in both groups, showing that transfer of collocational patterns from the L1 occurs half of the time.

Syntactic dissonance is found in all routes except perceived equivalents and cores. The most syntactic confusion derives from lack of suffix knowledge of base forms of words (associational dissonance), but a small percentage of syntactic dissonance stems from L1-interference, manifested in cognates and transliterations. Invalid, the smallest effect, is found in cognates, transliterations, synonyms and associations. Invalid dissonance can be divided into L1-based (concepts or phrases from the L1 that do not exist in the L2) and L2-based (non-words that are a result of lack of knowledge of spelling (associations) or compound words that do not have direct influence from the L1.

7.2.4 Influence

The subject of strong L1-influence has been discernible through the L1-based routes thus far in the data analysis. However, this does not explain interlingual and intralingual influence in the L2-based routes (L2-based routes can have weak L1-influence). No comparison has yet been done of how influences are distributed throughout effects either. As seen in table 5, possible L1-influence can be detected in 57% of the intermediate dissonances and 54% of the advanced dissonances in total. By examining the distribution, it may be possible to determine under what circumstances interlingual or intralingual influence is dominant. An investigation of distribution may indicate how/when students use and rely on their L1, and how/when they do not. In instances where the L1 is not a factor, suggestions for intralingual interference will be provided. This section will explore the distribution of inter- and intralingual influences in routes and effects. This will be done by showing distribution of influence in L2-based routes and effects, a log-likelihood test to determine if the differences between strong and weak L1-influence and intralingual influence are statistically significant between the two groups, and a comparative distribution to see if the proportion of distribution is the same or different between the two groups.

7.2.4.1 Routes

L1-influence in routes suggests to what extent Norwegian speakers of English depend on their L1 when they use words they have partial knowledge of. Table 12 shows how many inter- and intralingual influences were detected in routes. As explained in section 5.2.3, influences are subdivided into three types, where strong L1-influence represents L1-based routes (cognates and transliterations), and weak L1-influence and intralingual influence are both found in L2-based routes (perceived equivalents, synonyms, associations and cores):

Table 12. Distribution of influences in intermediate and advanced dissonances

	Intermediate students	Advanced students
Strong L1-influence	190	122
Weak L1-influence	244	78
Intralingual influence	270	173
Total	704	373

Table 12 suggests that L1-influences (both weak and strong) account for between 54-62% of all dissonances. Strong L1-influences account for between 27-32% of dissonances, meaning students depend directly on their L1 only around 30% of the time when choosing the wrong word.

7.2.4.1.1 Distribution of inter- and intralingual influence in L2-based routes

One phenomenon which table 12 does not explain is the distribution of inter- and intralingual influence in the L2-based routes. Whereas cognates and transliterations have 100% strong L1-influence, the inter/intralingual relationship is not apparent in perceived equivalents, synonyms, associations and cores.

Table 13. Inter- and intralingual influence in the L2-based routes

Route	Intermediate Students		Advanced Students	
	Interlingual	Intralingunal	Interlingual	Intralingual
PE	0	21	0	8
Synonym	13	117	27	112
Association	1	51	1	21
Core	230	91	50	110

Figure 25: Distribution of influence in L2-based routes for intermediate students

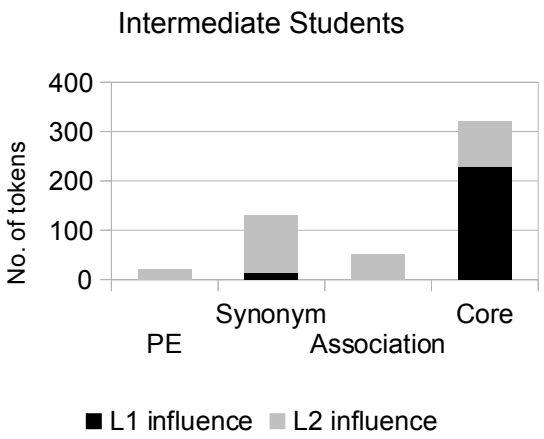
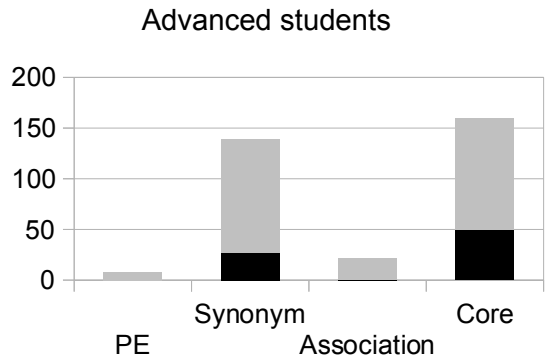


Figure 26: Distribution of influence in L2-based routes for intermediate students



As shown in table 13 and figures 25 and 26, synonyms are not influenced by L1 to a great extent. Only 10-24% of synonyms are influenced by finer gridding (see section 4.2.1). Cores show much more evidence of possible L1-influence, through the way core words in Norwegian and English are perceived (see section 4.2.1.2). Between 45-74% of core dissonances indicate possible L1-influence. Reasons for this could be the way core words are perceived in Norwegian and English, and how the norms for writing formal texts is different between the two languages.

To conclude, strong L1-influences claim between 27-32% of all dissonances. L2-based routes show traces of finer gridding and possible stylistic influence, and these weak L1-influences account for between 21% (advanced) and 35% (intermediate) of all dissonances. The remaining percentages of dissonance (38- 46%) are words that have possible intralingual (L2) influence.

7.2.4.1.2 Statistical significance

Influences were tested for statistical significance in strong and weak L1-influence, and intralingual influence. The log-likelihood calculator gave the following results.

Table 14. LL-values and p-values for statistical differences between intermediate and advanced influences

Influence	LL-value	P-value	Result
Weak L1	86.93	P < 0.0001	√
Intralingual	19.73	P < 0.0001	√
Strong L1	13.76	P < 0.001	√

√ = difference between intermediate and advanced is statistically significant

x = difference between intermediate and advanced is not statistically significant

The results of the log-likelihood test suggest that there is a statistical difference in all types of influence. Weak L1-influence gives the largest LL-value, suggesting that this is where intermediate and advanced students differ the most. Intralingual and strong L1-influence have much lower LL-values, reflecting that the difference in these two categories is smaller between intermediate and advanced students, yet still significantly different.

7.2.4.1.3 Comparative distribution

Figure 27: Comparative distribution of influence in intermediate students

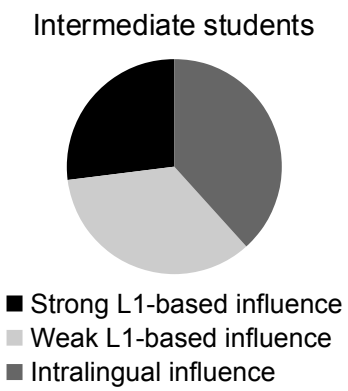
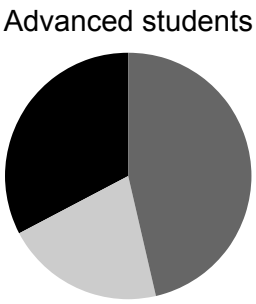


Figure 28: Comparative distribution of influence in advanced students



The comparative distribution of influences shows that influence is proportionally different in the two groups. Weak L1 influence is much more preferred by the intermediate students, while advanced students rely on strong L1 influence and intralingual influence. We will return to potential reasons for this in the next section.

7.2.4.2 Effects

Examining influence in effects suggests how much Norwegian speakers of English depend on their L1 when they have insufficient knowledge of an L2 item. An examination of the distribution of L1-influence may give insight into which aspects of word knowledge Norwegian speakers of English perceive as being more "transferable" from the L1 to the L2 and which aspects of word knowledge they would rather turn to their L2 for. For instance, since stylistic dissonances have a high percentage of L1 influence, this may indicate that Norwegian speakers of English perceive style as being similar between Norwegian and English, and thus acceptable to transfer. Alternatively, the gap of stylistic knowledge in Norwegian learners of English is much larger than other aspects of word knowledge, and the way Norwegian learners of English cope with this gap is by L1-transfer.

Table 15. Total number of influences found in effects for intermediate students

Effects	Intermediate Students		
	Strong L1-influence	Weak L1-influence	Intralingual influence
Semantic	57	1	59
Collocation	44	10	34
Styl/Con	60	232	140
Syntactic	8	0	28
Invalid	21	1	9
Total	190	244	270

Table 16. Total number of influences found in effects for advanced students

Effects	Advanced Students		
	Strong L1-influence	Weak L1-influence	Intralingual influence
Semantic	20	3	28
Collocation	27	8	19
Styl/Con	63	67	110
Syntactic	8	0	13
Invalid	4	0	3
Total	122	78	173

7.2.4.1.2 Distribution of influences in effects

Figure 29: Distribution of influence in effects for intermediate students

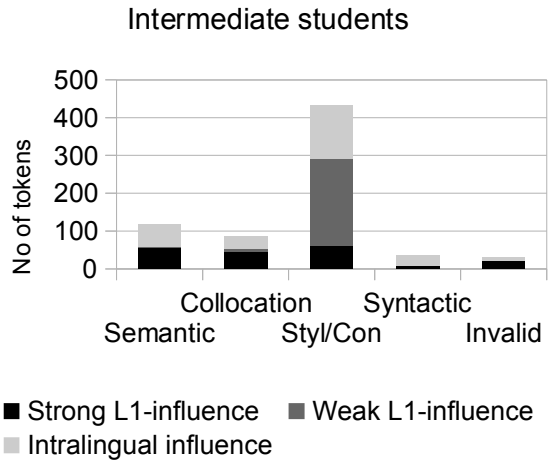
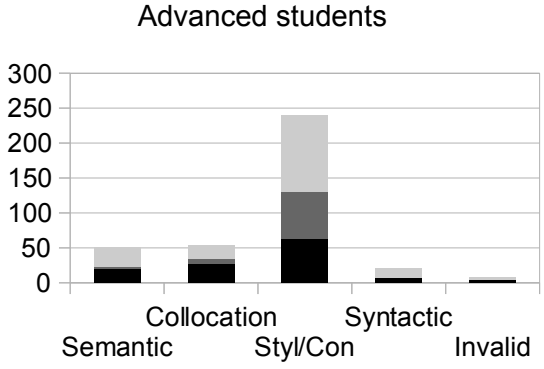


Figure 30: Distribution of influence in effects for advanced students



The largest difference in distribution of influences in effects appears to be stylistic/connotational, where intermediate students seem to rely more on equivalent L1 core words (*big* - *stor*), and advanced students prefer to use an L2 word that a) cannot be explained by any perceivable L1 influence (e.g. ***a bunch of***), or b) is a directly translatable non-core word (*totally* - *helt*). Stylistic and collocational are the only effects that show weak L1-influence, meaning that these effects are the result of weak finer gridding and the core word stylistic differences.

Distribution of semantic dissonance suggests that students will either refer to a directly translatable word for sense meaning, or there is an unknown intralingual influence that causes the student to make an incorrect hypothesis of what a word means. This suggests that weak finer gridding does not occur on a semantic level – perhaps because words that are synonyms are more likely to have different collocational and stylistic patterns than cover a different semantic space. Syntactic dissonance is affected much more by intralingual influence, indicating that confusion of word forms usually stays within the L2.

In both learner groups it is clear that stylistic/connotational dissonance has the most L1-influence. For the intermediate students, 68% of all stylistic/connotational dissonance suggest a strong or weak L1 influence, and for the advanced students, this influence is as high as 54%. This would suggest that the highest number of unsuccessful transfers from Norwegian to English is stylistic. In other words, Norwegian speakers repeatedly use English words with a stylistically different Norwegian equivalent (*big* versus *stor*, for instance). This accounts for 63% of all stylistic dissonance for Norwegian speakers of English. The source of the problem is then that English equivalents of Norwegian words do <u>not</u> have the same stylistic appropriateness, since Norwegian words that can work on both a formal and an informal level, can decidedly <u>not</u> work on a formal level in English. Stylistic dissonance is found in both L1 and L2-based routes. 21% (intermediate) and 48% (advanced) of stylistic dissonances are L1-based. This indicates that advanced students transfer much more stylistic knowledge directly from their L1, whereas intermediate students have more indirect L1-influences, such as the case of the different behavior of core words in Norwegian and English.

One must keep in mind that 37% of all stylistic dissonances are not L1-influenced. This can be explained by the students' preference for core words and colloquial words (synonyms). This type of knowledge is L1-independent, meaning that L2 preference for informal words is an intralingual factor.

The second largest L1-influenced effect is collocation. 61% (intermediate) and 65%

(advanced) of all collocations have L1-influence, implying that in over half of the cases, when a student is uncertain of how to combine words (***big*** *demand* (*high*), ***be*** *a part* (*play*) or *plane* ***accident*** (*crash*)), there is a definite L1 word combination equivalent which may have influenced the speaker's choice of word combination. Nearly all of these collocational L1-influenced dissonances are strong, accounting for as much as 77% (advanced) to 81% (intermediate) of all collocational dissonance. The rest of L1-influenced collocations are the result of weak finer gridding, and only a few examples can be found in synonyms and associations.

The remaining non-L1-influenced collocations, which are mainly synonyms and associations, account for as much as 39% (intermediate) and 35% (advanced) of collocational dissonance, suggesting that in slightly less than half of the cases of word combination dissonances, there is no apparent evidence of an L1 "crutch." There are several possible intralingual reasons why these word combinations are not the result of L1-influence: the L2 speaker may have not acquired the correct synonym, has chosen a synonym s/he is more comfortable with, used the only synonym s/he knows, or chose the wrong of two synonyms (*a smile crossed her* ***mouth*** (*lips*)). Choosing the wrong synonym can be an arbitrary choice or a confusion of word forms (associational dissonance), such as: *when words* ***crash*** (*collide*).

The third largest L1-influenced effect is semantic, with between 45% (advanced) and 50% (intermediate) of all semantic dissonance suggesting L1-influence. As in the case of collocations, most semantic dissonance with L1-influence stems from cognates and transliteration (57 out of 58 tokens for intermediate, 20 out of 23 tokens for advanced). Cases where L1-influence is not apparent in semantic misunderstanding include perceived equivalents, where the student uses a word without knowing its meaning (e.g. *give a* ***withering*** *look*), synonyms where semantically similar words confuse the L2 speaker (***divisions*** *of a hospital* instead of *wards*) and associations, where similar word forms that have unrelated meaning will cause the L2 speaker to choose the wrong word (*wrong face* ***prepositions*** instead of *proportions*). Semantic dissonance in perceived equivalents, synonyms and associational routes are therefore primarily intralingual, illustrating that non-L1-influenced semantic dissonances stems from confusing the right word with a word that is semantically similar (synonyms), similar in word form (associations), or has no relation to the correct word at all (perceived equivalents). These intralingual factors account for roughly half of the semantic dissonance whether the student is at an intermediate or an advanced level.

7.2.4.3 Summary of influences

In conclusion, the three largest L1-influenced effects are style/connotation, collocation and semantic. Stylistic influence is the result of both L1- and L2-based routes, with more dissonance in the L2-based routes. Intralingual factors suggest preference for use of more informal words (cores and synonyms). Collocational and semantic dissonance are primarily the result of L1-based routes, and therefore show strong L1-influence. Intralingual factors for collocational dissonance include not having acquired the right synonym, preference for the wrong synonym, arbitrary guesswork (synonym), and confusion of word forms (associations). Intralingual factors for semantic dissonance include picking a word that is semantically similar to the right word (synonym), confusing two words that are similar in form (association), or unknown factors that are highly individual (perceived equivalents).

The log-likelihood test shows that the difference between the two learners groups is statistically significant in all three influences; all three influences are significantly larger in the intermediate group. The comparative distribution informs us that the distribution within each group is different: intermediate students depend on weak L1-influence, and advanced students prefer strong L1-influence and intralingual influence.

8. DISCUSSION

This section will revisit the research questions presented in chapter 2. The discussion will first look at the first research question: how lexical dissonance is generally distributed in Norwegian speakers of English. Subquestion a-e will be examined in detail. The second research question, which compares distribution of intermediate and advanced students, will be examined afterwards, by answering subquestions a-e.

8.1. The first research question

1. How is lexical dissonance manifested in free production texts written by Norwegian speakers of English (intermediate and advanced)?

a) *How is lexical dissonance distributed across routes and effects in Norwegian intermediate and advanced learners of English?*

Lexical dissonance is classified through both routes and effects. For routes, the distribution in descending order is: cores, synonyms, transliterations, cognates, associations and perceived equivalents. For effects, the distribution in descending order is: stylistic/connotational, semantic, collocational, syntactic and invalid. For a full overview of distribution, see figures 5-7.

b) *If some routes and effects are larger than others, what are possible reasons for this?*

Routes

Routes are the way a speaker chooses the wrong word, and reflects some of the psychological processes behind "choosing." For instance, will a learner prefer to pick a word similar to the L1 (cognates, transliterations) or a simplified term (cores)? Some routes are clearly larger than others in this study of lexical dissonance. There seem to be two contributing factors:

i. **Some routes are "wider" than others.** Some routes cover a larger range of types of lexical dissonance than others. In the case of cognates and transliterations, there are fewer words between Norwegian and English that are similar in form and meaning (*high* - *høy*) than similar in meaning and not in form (*fin* - *nice*). Therefore, there are more opportunities for

transliterational dissonance, since the number of cognates are more limited. The same can be said of synonyms and associations – there are fewer words that are similar in form and meaning (*increase - enhance*) than words that are similar in meaning and not form (*seemingly - apparently*). Since synonyms and transliterations are "wider" routes, they rank much higher in types than cognates and associations.

ii. **The route is preferred by the L2 speakers**. The allowance for more types of dissonance in some routes, as explained above, cannot explain why, for instance, there are so many dissonant cognates. Even though some routes will allow for many different types of dissonance does not mean that speakers will necessarily use the route. Some routes are simply preferred to others.

A closer look at each route

Cores is the the most frequently chosen route, representing 4 in 10 wrong words. This overwhelming preference would suggest that Norwegian speakers of English use core vocabulary in cases where L1 English speakers would use more precise, low-frequency words, and this occurs often in argumentative texts. This is seen by Schmitt as being one of the most common mistakes in L2 learners (Schmitt 2000: 155). Norwegian speakers of English satisfy their lexical needs by using basic vocabulary instead of more specific, appropriate terms (using *big* instead of *large*, *tremendous* or *incredible*), and this will often lead to production-blocking (there were 20 instances of *I think* instead of *I believe/am of the opinion*).

Synonyms is the second largest route and accounts for 1 in every 5 wrong words. This route shows that Norwegian speakers of English have a range of words available in their vocabulary, but their knowledge of these words is partial, and choosing between similar words is a (seemingly) arbitrary decision. Dissonance then occurs when L2 speakers choose the wrong synonym, either by believing synonyms are interchangeable, or by choosing the synonym they know/prefer. Since this is a popular route, particularly among the advanced students, this would suggest that choosing the right synonym can be a challenge and a large source of dissonance.

The numerous instances of **transliterations** and **cognates** (1 in every 6 wrong words for both routes) illustrates that many Norwegian speakers of English perceive Norwegian and

English as having much more in common than they do in reality. Many of the dissonances found in these two routes support the word equivalent hypothesis, i.e. that Norwegian speakers of English habitually transfer all aspects of word knowledge from their L1 to their L2, assuming that the two will have a 1:1 mapping. The strong relationship between semantic dissonance and cognates suggest that speakers will often assume that cognates have the same meaning in the L1 and L2, and the strong relationship between transliterations and stylistic dissonance shows that speakers perceive translatable words as having the same stylistic markings.

The low number of **associational dissonances** (1 in every 14 dissonances) versus the high number of synonymous dissonances suggests that Norwegian speakers of English do not rely primarily on phonological routes, i.e. that there is more dependence on the meaning similarity between words than the sound similarity. The small percentage of **perceived equivalents** (1 in every 33 dissonances) suggests that Norwegian speakers of English either a) are reluctant to use words they are very uncertain of or b) have some idea of the words in their vocabulary (and will therefore not make perceived equivalence dissonance).

Effects

Effects reflect the "incorrect hypotheses" an L2 speaker creates when s/he decides to use a word s/he only possesses partial knowledge of, i.e. the assumptions speakers naturally make when they do not fully know a word. The reason why some effects are larger than others can be due to:

i. **The speaker lacks some aspects of word knowledge more than others**. For instance, stylistic dissonance is very frequent. This could be explained by either a) stylistic dissonance is not covered sufficiently in the EFL classroom or b) stylistic knowledge is harder to acquire than other types of word knowledge (Schmitt 2010: 31).

ii. **The speaker perceives some aspects of word knowledge as being more important than others.** If speakers perceive stylistic knowledge as being less important, they may be more willing to risk using words they have no stylistic information about. Conversely, if they perceive semantic knowledge as being important, they will not risk using a word they have no semantic information about, causing more stylistic dissonance and less semantic dissonance. This does not necessarily mean that they have more knowledge of other word aspects, only

that they are more willing to be wrong in some aspects and not others.

iii. **Some aspects of L1 words are viewed as being more transferable than others**. In the case of effects that have high L1-influence, incorrect hypotheses could be made because the student makes a false 1:1 mapping between an L1 and L2 item. These are instances of false word equivalences (see section 4.2.1.1) and may explain why students are more willing to use words they only have partial knowledge of.

A closer look at each effect

Stylistic/connotational dissonance constitutes the largest effect at over 1 in 2 wrong words. This effect mostly consists of dissonance between informal words in formal contexts. Influence distribution in effects suggests that this is a large effect because a) students are often influenced by equivalent L1 items that are stylistically neutral, meaning Norwegian speakers of English perceives L1 words as being stylistically transferable and b) Norwegian speakers of English have acquired many stylistically informal words that have no perceivable L1-influence. This suggests Norwegian speakers of English acquire English through colloquial sources, such as television and the internet. Norwegian speakers of English do not produce the "mishmash of styles," which Martin claims is "characteristic of non-native production," – they stay true to one register (informal), but this seems to be the only register they know (Martin 1984: 131).

Semantic dissonance is the second largest effect, accounting for 1 in every 6 dissonances. The distribution of influences in effects suggests that in half of the cases, words are semantically dissonant because of the semantic equivalence hypothesis, i.e. the student will transfer semantic features of an L1 word to a similar L2 word. However, the other half of semantic dissonances suggest intralingual influence; the difficulty lies in choosing which synonym is "correct." Meaning is often focused on in the EFL classroom, and most would undoubtably agree that it is one of the more important aspects of word knowledge, if not the most. Therefore, semantic dissonance can reflect that meaning is complex and difficult to acquire.

Collocational dissonance is the third largest effect, and also at approximately 1 in every 6 dissonances. Similar to semantic effect, collocational knowledge is transferred from the L1 to the L2 in approximately half of the wrongly chosen words. A small proportion of

dissonant collocates are caused by weak finer gridding, and the rest are due to intralingual influence. This could mean that when a Norwegian speaker of English is unsure of how to combine words (particularly which verb will combine with which object), half of the time they will refer directly to L1 knowledge, the other half they will choose a synonym they know/prefer, and in some select cases this will be the wrong of two synonyms for which their L1 only has one word to express (weak finer gridding). Schmitt suggests that collocational knowledge is difficult to acquire because it calls for vast amounts of exposure to the language (Schmitt 2010: 31), so this could explain why students struggle with this aspect.

Syntactic dissonance is the second smallest effect, and only counts for 1 in every 20 dissonances. Mostly it is caused by intralingual influence due to confusion of similar words within a word family, such as a ***beautifully*** *day*. There are a few select instances where syntactic elements of an L1 word will be transferred, such as ***man****-dominated* from Norwegian *mannsdominert* (*male-dominated*). The many cases of intralingual influence suggest that syntactic knowledge is perceived as less transferable from the L1 to the L2, as opposed to the other effects. The low number of syntactic dissonance either suggests that this aspect of word knowledge is less difficult to acquire, the EFL classroom focuses on syntactic elements more, or students are hesitant to use words if they are unsure how to place them syntactically in a sentence.

Invalid dissonance is the smallest effect and is also approximately 1 in every 20 wrong words. It accounts for words or structures that do no exist in the L2, and is largely the result of creating new words and concepts based on the L1, e.g. ***psychical*** *damage* (*psychological damage*). Invalid (4%) is primarily the transfer of Norwegian concepts/words to English (***cold degrees*** from Norwegian *kuldegrader*) (although it also covers non-existent words caused by intralingual influences). Although transfer is common in the other effects, there were only 38 instances of invalid structures in a body of 50, 000 words. This could imply two things:

a) Norwegian speakers are aware of which concepts can and cannot be transferred from Norwegian to English. If this is the case, this would also imply that Norwegians have a high awareness of the contrast between Norwegian and English concepts and culture, since many invalid structures (*life lie* - *livsløgn*) are a reflection of Norwegian culture and how Norwegians perceive the world.

b) Transfer is practiced frequently, but since so many concepts and words are similar between Norwegian and English, it is rare to find instances where a Norwegian structure has

no English equivalent. A performance analysis of transfer (calculating how many positive transfers there are versus negative) may be able to confirm or refute such a statement.

c) *How do routes and effects combine? Is there a discernible pattern between how a wrong word is chosen and why it is dissonant?*

It was found that some route/effect combinations are more frequent than others.[11] **Cores** are nearly all stylistically dissonant, because it is rare for core words to not collocate with surrounding words (***good*** *guy/time/food*), not work syntactically or be misunderstood. In other words, the only way a core word can be dissonant is when native speakers would have used more precise vocabulary in a formal register.

Cognates, **transliterations**, **synonyms** and **associations** all have a more even spread of effects. This would suggest that words in these routes can be misunderstood on all levels. Transliterations and synonyms have more stylistic dissonance, suggesting perhaps that cognates and associations are relied on less for stylistic information. "Associations" has a large proportion of syntactic dissonances due to that words from the same word family placed incorrectly in a sentence (*beautiful* - *beauty*) are labeled as associational mistakes. Cognates and transliterations have a high number of invalid dissonances; this is because "invalid" caters mostly to structures that exist in Norwegian and not English (e.g. *life lie* - *livsløgn*), and these structures can only be found in L1-based routes.

d) *What is the proportion of L1-influenced dissonances (interlingual) vs L2-dissonances (intralingual)?*

In this study, 59% of all dissonances had perceivable L1-influence, suggesting that as much as 41% of dissonances can be explained by intralingual factors. Ellis writes that most learner errors are intralingual in origin (Ellis 2008: 55). However, the findings from this study suggest that a majority of lexical errors are interlingual in nature.

11 For a full overview of the distribution, see tables 10 and 11

e) *Will my data have the same distributional patterns for routes and effects as Hasselgren's?*

Route distributions

Hasselgren: **synonyms > transliterations > associations > cognates = PE > cores**

This study: **cores > synonyms > transliterations > cognates > associations > PE**

Effect distributions

Hasselgren: **semantic > collocational > stylistic/connotational > invalid > syntactic**

This study: **stylistic/connotational > semantic > collocational > syntactic > invalid**

The distributional patterns are very different in both routes and effects. The most striking difference is cores, which ranks highest in my distribution and lowest in Hasselgren's. Synonyms and transliterations rank high for both studies, while associations, cognates and perceived equivalents rank on the lower end of the scale for both studies.

The other large difference is stylistic/connotational, which is by far the largest effect in this study, and only the third largest effect in Hasselgren's study. Semantic effect is prominent in both groups, and invalid/syntactic dissonance is low.

There are many reasons for the different results. Primarily the type of data can explain this, as translation data will elicit different types of dissonance than free production texts. Another possible explanation is that Hasselgren and I label dissonance differently – I might have considered more core words as “dissonant”, or she might have viewed these dissonances as being synonymous. A third possibility is other variables – the participants in our groups are from different time periods and have had different exposure to English, English education in Norwegian has changed from 1993-2012, or some results could be due to chance.

8.2 The second research question

2. Will the number and distribution of lexical dissonances differ between the intermediate and advanced students?

a) *Is there a difference in number of lexical dissonances found between intermediate and advanced students?*

(See table 4).

The intermediate students have 331 more dissonances than the advanced students, i.e. number of dissonances diminishes by 47% in the advanced texts. This would confirm that there is a clear difference in number of lexical dissonances for intermediate and advanced student texts.

b) *Is there a difference in distribution of lexical dissonance between routes?*

(See table 6 and figures 12-14).

The comparative distributions suggest that there is a difference in distribution of routes between the intermediate and advanced students, even if we disregard difference in size of the routes. It appears that cores are the preferred route of intermediate students, while advanced students' routes are more evenly distributed between synonyms, cores and transliterations. The routes cognates, perceived equivalents and associations do not show any significant changes in distribution. These results suggest that advanced students have different ways of choosing wrong words compared to the intermediate students, and prefer a) different synonyms and b) structures that only work in the L1 (transliterations) when picking a word, whereas intermediate students will pick “safe” core words.

c) *Is there a difference in distribution of lexical dissonance between effects?*

(See table 8 and figures 18-20)

Whereas there is a difference in distribution of routes, there is no perceivable difference in distribution of effects. This would suggest that although Norwegian speakers of intermediate

and advanced English have different ways of choosing the wrong word, the effect of these wrong words remains the same in both groups.

d) *If there are differences between intermediate and advanced students, are these differences statistically significant?*

The answer is both yes and no. The difference between number of dissonances overall is **p < 0.0001,** suggesting that there is a very significant difference between the two groups. When it comes to routes, cores, cognates, associations and perceived equivalents showed positive for statistical significance, while synonyms and transliterations did not (see table 7). All effects but syntactic showed positive for statistical significance (see table 9). To conclude, overall there is a significant difference between the two groups, but the distribution of routes and effects is not statistically significant in all cases.

e) *Will the advanced students show proportionally less L1-influence in lexical dissonance than the intermediate students, illustrating a "reliance on their ever-increasing knowledge of the target language" as suggested by Taylor (1975)?*

(see table 12 and figures 27-28).

The intermediate students show evidence of L1-influence in 62% of all dissonance, with a lower proportion of the L1-influences showing strong influence, and a higher proportion of them showing weak influence. The advanced students show evidence of L1-influence in 54% of all dissonances, with a higher proportion of the L1-influences showing strong L1-influence, and a lower proportion of them showing weak influence. One can conclude that proportionally speaking, advanced students show marginally less L1-influence (8%), but there is a preference for strong L1-influence. This can be explained by the high number of transliterations found in the advanced student dissonances and the high number of cores found in the intermediate student dissonances. The fact that advanced speakers prefer transliterations proportionally more than intermediate students may suggest that advanced speakers rely more on the knowledge of their first language when they wish to express themselves in ways they have not learned in their L2. Intermediate speakers, on the other hand, will stick to "safe" core words, such as *big*, or *nice* that have equivalents in their L1 (*stor, snill*). However, the advanced students also show more intralingual influence than intermediate students (44%

advanced and 38% intermediate). A possible explanation for this is while intermediate students again rely on "safe" core words, advanced speakers are more prone to using a wide array of synonyms that they only have partial knowledge of. This shows signs of "reliance in knowledge of target language" (Taylor 1975: 88). To conclude, advanced speakers show signs of relying more heavily on strong L1-influence and intralingual influence, while intermediate speakers tend to rely on basic words that often have an equivalent in their L1 (*big - stor*).

Interlingual influence should be measured along a continuum, ranging from direct L1-influence (*how goes it - hvordan går de*t), to possible instances of L1-influence, such as weak finer gridding (*the ideal family* ***female*** instead of <u>*woman*</u> (*kvinne*)). In certain stages of interlanguage (beginner, intermediate, advanced), L2 speakers rely on different ends of the continuum (from strong to weak). Taylor's statement suggests that progressing from a beginner to an advanced L2 stage will mean a parallel progress on the interlingual influence continuum from strong to weak (i.e. beginner L2 speakers will rely on strong interlingual influence, while advanced students will rely on weak L1 influence and gradually intralingual influence). This, however, is a fallacy, as advanced students show a preference for strong influence. Conversely, many intermediate students are more willing to rely on core words (weak L1-influence) and "sacrifice" some of the nuances that come with possible positive transfer.

9. CONCLUSION

This study has investigated the distributional patterns of lexical dissonance in texts written by Norwegian learners of English. The distributional patterns were determined based on a taxonomy of lexical errors developed by Angela Hasselgren with Norwegian speakers of English in mind. The taxonomy was altered to fit free production data and categorized lexical dissonances based on their source (inter/intralingual influence), how L2 speakers chose the wrong word (routes) and why the word was wrong (effect).

The main research question was: how is lexical dissonance distributed in texts written by Norwegian speakers of English? A quantitative analysis determined that for **routes** the distribution was: cores (39%), synonyms (22%), transliterations (16%), cognates (13%) associations (7%) and perceived equivalents (3%). For **effects**, the distribution was: stylistic/connotational (62%), semantic (16%), collocational (13%), syntactic (5%) and invalid (4%). For **influences**, the distribution was: L1-influence (58%)[12] and intralingual influence (42%). Furthermore, it was observed that positive correlations existed between the relationship of some of the routes, effects and influences.

The second research question stated: will the number of lexical dissonances and their distributions differ between intermediate and advanced texts? A log-likelihood test established statistically significant differences in the number of overall dissonances. These significant differences indicate that intermediate students have significantly more lexical dissonances than advanced students, which then implies that lexical dissonance is more commonly found in intermediate speakers of English. In **routes**, cores, cognates, associations and perceived equivalents were significantly different between the intermediate and advanced students, while transliterations and synonyms were not. In **effects**, semantic, collocational, stylistic/connotational and invalid dissonance were significantly different, but syntactic dissonance was not. In influences, all categories (strong L1-influence, weak L1-influence and intralingual influence) were significantly different.

Comparative distributions suggested that the way students pick the wrong word (routes) is different between intermediate and advanced students, but how these words are wrong (effects) is the same. Influences also suggested that advanced students rely more on direct L1-knowledge when they have partial knowledge of a word, and intralingual influence is also a large cause of dissonance. Intermediate students are influenced primarily indirectly

12 Strong L1-influence (29%) and weak L1-influence (29%).

by the L1 (weak L- influence), and this is largely due to picking inappropriate L2 words that are appropriate in the L1.

9.1 Pedagogical implications

Llach stresses that in light of results obtained from an examination of lexical errors, it is the task of the researcher to "deduce and extract the behavioral and developmental patterns of lexical errors, and design pedagogical solutions to the lexical problems evidenced by the analysis of lexical errors" (Llach 2011: 91). This thesis has explained some of the "behavioral and developmental" patterns of lexical dissonance in Norwegian speakers of English, with results that strongly indicate that patterns of lexical dissonance exist. It should therefore be possible to create solutions for lexical dissonance that can be implemented in the Norwegian EFL classroom.

A problem with designing pedagogical solutions from the findings of such a lexical error study is that it does not take certain factors into account. This study and taxonomy are mainly **descriptive** in nature – they explain the types of lexical dissonance, how they are chosen (routes) and possible influences, but cannot explain fully why some effects are larger than others. This is caused by other factors, such as:

i. The priorities and goals of the EFL classroom

ii. The L2 speaker's exposure to English and linguistic environment

iii. Developmental patterns that follow a certain sequence of progression

The goals and focus of the Norwegian EFL classroom may explain some of the results – for instance, if semantic accurateness is more important than register-training to the Norwegian educational system, Norwegian students of English will then learn and be more aware of semantic aspects of English. This can then account for why semantic dissonance is so low and stylistic dissonance is so high. Additionally, even if a student has, say, more semantic dissonance, s/he may be more wary of this in the EFL classroom because it is a focus point. The L2 student may therefore consciously use a dictionary or other resources when given the chance, reducing the number of naturally occurring dissonances.

The L2 speaker's linguistic environment is also a factor – recent trends in Norwegian youth television habits suggest a large exposure to colloquial American English, and may explain why students are mainly aware of informal register, and lack the vocabulary of a

formal register. Lastly, developmental patterns of lexical knowledge (which there is little information on, as linguists still cannot agree on what lexical knowledge fully consists of) may be the cause of some incorrect hypotheses. In other words, it is possible that stylistic knowledge "lags" behind in the mental lexicon naturally, and cannot be explicitly taught.

Due to these variables, this study can only account for some of the largest areas of lexical dissonance, and how these areas could be improved upon. In very brief terms, dissonance reveals that the following areas of lexis are most problematic for Norwegian intermediate and advanced speakers of English:

i. Lack of knowledge of formal/specific words

ii. Overuse of basic vocabulary

iii. Lack of stylistic knowledge of words

iv. Dependence on L1 for word knowledge

Points 1, 2 and 3 overlap. Because there is a lack of specific vocabulary, L2 speakers compensate by overusing their familiar, basic vocabulary, and this basic vocabulary is often too informal. However, point 3 is also independent – stylistically informal words are only "core" words between 45-72% of the time. This means that 22-55% of all stylistic dissonances are **not** caused by overuse of basic vocabulary, but rather the result of the learner's exposure to informal, colloquial language.

The obvious solution to point 3 would be to expose Norwegian students of English to more formal/academic English; implementing such a solution in the EFL classroom is an entirely different matter. Norwegian students of English surely read formal English in class, but the amount needed to outweigh the informal English input they receive from television and internet could simply not be covered in an EFL classroom. Therefore, strategies to elicit stylistic knowledge and build awareness could be a possible solution. An example would be for a teacher to write several similar words on the board and have students categorize them based on their formal/informal features. Training students to be aware of stylistic differences may aid them into independently categorizing words based on their stylistic markings in the future. However, this does not solve the problem of **lack** of formal/specific words (point 1). A solution could be to boost vocabulary, but what type of specific/formal words do the students need to acquire? If the topic is "slavery in America," the students have no problem using specific words and concepts that are associated with that time period, such as *segregation,*

indentured servants and *white power*.

The words students tend to overuse are in most cases verbs (for instance, the many cases of *get*), adjectives (*big, good, bad*) and adverbs (*really, very*). It is difficult to explicitly teach students how to use more precise words in these cases, and would be perhaps easier to illustrate which words are "core" and should be avoided in formal writing. For instance, if a teacher is aware that several students in the case are overfond of using *I think*, they can offer the solutions *I believe/am of the opinion*. The important part is teacher awareness of stylistically informal words. Appendices 3 and 4 provide lists of core words which could be a good starting point for basic words that Norwegian students of English generally struggle with and need alternatives for.

Point 4 poses a problem because transfer is a very common and natural phenomenon, and does not always lead to dissonance. Therefore, it is not something that can or should be unlearned. Here one should differentiate between strong and weak L1-influence. The L1-based routes are perhaps more conscious and obvious instances of transfer, and can be dealt with in the EFL classroom more easily. Weak L1-influence (as in the case of core words) is often done subconsciously, and is much more difficult to pinpoint. This discussion will refer to strong or L1-based influences.

False cognates form a relatively small group and can be easily categorized and explained. Appendices 3 and 4 provide lists of some of the most common false friends Norwegian students of English struggle with (but it is by no means exhaustive). If teachers were to create such a list of false friends, perhaps after observing their own students' texts, they could then explain some of the main differences between Norwegian and English cognates to build student awareness that not all similar words have a 1:1 mapping. Transliterations are trickier. As suggested in section 8.1, transliteration forms a much larger group of types of dissonance. Students frequently transfer phrases, word combinations and idioms with and without success. This habit is in some cases viewed as positive because Norwegian and English are similar enough for some transfers to be successful. If a student is unsure of a word, transfer is a useful resource. However, automatically assuming that all Norwegian words and English equivalents have a 1:1 mapping will inevitably lead to many dissonances with strong L1-influence. Therefore, students should be encouraged to use transfer only as a last resort strategy – if the student knows another, simpler way of expressing her/himself, it is better to use a "correct" and simpler L2 item than "risking" a transliterational failure. The best way to prevent negative transfer in the EFL classroom is to stress that

Norwegian and English are not as similar as we may like to believe, and therefore not all transfers will be successful. This can be illustrated by handing out lists of Norwegian phrases and collocations that have no or a different English equivalent.

9.2 Suggestions for further research

There is of course much more that could have been done with the data given, but due to time and space constraints (and a range of other factors), these aspects could not be investigated in this thesis. This section includes suggestions for further work that could have been done with the data in this study, e.g. eliciting more information through different/more methods, and further work that could be done by gathering/using different types of data.

Further investigation on the present data:

- **Using performance analysis to examine the data.** Performance analysis can be useful to determine how successful versus unsuccessful the students are, e.g. comparing how many correct and incorrect collocations are found in the texts. This could give an overall performance evaluation of what students struggle and are successful with in their lexical production.

- **Using another taxonomy of lexical errors to examine the data**. Different taxonomies reveal different types of errors. For instance, Hemchua & Schmitt (2006) created a taxonomy with 24 categories of lexical errors for a fine-grained analysis (as cited in Llach 2011: 77). Hemchua and Schmitt's taxonomy also includes form-oriented lexical errors, which were not examined to a great extent in this study. Using such a taxonomy may reveal different aspects of lexical error.

- **Investigation of word classes in lexical dissonance.** Classifying dissonances based on their word class may illustrate which word classes the students struggle with the most. For instance, what collocational combinations do students find the hardest? Those with verbs, or those with nouns?

- **Grammatical errors and lexical dissonance comparison**. A study of grammatical errors and lexical errors may indicate if the two correlate. For example, will an L2 speaker who produces many lexical dissonances also produce many grammatical errors?

Further work on additional data:

- **Longitudinal study.** If the intermediate students wrote similar texts in, for example, five years, this could show if there have been any patterns of development in lexical dissonance. For this data it is impossible, since the intermediate students were granted full anonymity when collecting their texts. However, longitudinal studies of lexical errors are uncommon, and the field would undoubtedly benefit from from such a study.

- **L1 and L2 comparisons.** By collecting data written in the students' L1 (Norwegian), it can be determined if the students who struggle with lexical dissonance in their L2 also struggle with lexis in their L1. Then a possible link can be drawn between L1 and L2 lexis. Again, for this particular data set it is not possible, since there is no access to information concerning either the intermediate or advanced students.

- **L1 English control group.** By comparing the student texts to similar texts written by an L1 English control group, it can be determined whether some of these dissonances are common regardless of whether English is L1 or L2. This is particularly useful for stylistic dissonance. Students were marked as "dissonant" for using basic, non-specific and informal vocabulary. However, have L1 English students through ages 15-21 acquired a specific, formal vocabulary? A comparison of L1 English texts could show if L1 students also rely on a colloquial, non-specific vocabulary when writing argumentative texts, and determine if this is generally a problem, or only a characteristic of Norwegian intermediate and advanced English. The ICLE corpus has similar texts written by English native speakers, and a comparison can be done between the Norwegian advanced and L1 English speakers.

- **A replica analysis for comparable results**. This study is the only study to my knowledge that examines distribution of lexical errors in free production texts written by Norwegian speakers of English. A replica study would therefore be useful to compare results. If a replica study had more data (i.e. more texts and participants), the results would be more robust and could strengthen or counter my results.

9.3 Final words

This thesis has hopefully demonstrated the importance of lexical words in the second language vocabulary. The analysis and results of this study suggest that lexical errors occur frequently and can impede understanding. By identifying the most frequent errors and developing systems for them, the second language teacher is better equipped to deal with lexical issues by knowing a) how and why lexical words are wrong, b) which types of lexical errors to anticipate and c) how to inform L2 learners so that these errors can be prevented.

This study can serve as a basis for anyone in the future who wishes to examine lexical dissonance further, whether it be Norwegian L1 or any other L1. The taxonomy has been altered to fit free production data which can allow for larger quantities of data, e.g. national tests. The lists of types of lexical dissonances located in appendices 3 and 4 can also be used by curious English teachers in Norway if they wish to know what types of mistakes are most common for Norwegian students of English.

10. REFERENCE LIST

Aitchison, Jean. (2012). *Words in the Mind: An Introduction to the Mental Lexicon*. UK: John Wiley & Sons, Inc

Carter, Ronald. (2012). *Vocabulary: Applied Linguistic Perspectives*. USA & Canada: Routledge

Carter, Ronald. McCarthy, Michael. (1988). *Vocabulary and Language Teaching*. NY: Longman Group UK Limited

Caspi, Tal. Lowie, Wander. (2010). "A Dynamic Perspective on L2 Lexical Development in Academic English;" in Chacón-Beltràn, Rubén; Abello-Contesse Christián; Torreblanca-López María del Mar. *Insights into Non-native Vocabulary Teaching and Learning*: 41-61.

Crystal, David. (2008). *A Dictionary of Linguistics and Phonetics*. USA; UK: Australia: Blackwell Publishing.

Cruse, D. A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.

Ellis, Rod. (2008). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Erman, Britt. Warren, Beatrice. (2000). "The Idiom principle and the open choice principle" in *Text. 20*, 20 (1): 29-62.

Fiedler, Sabine. (2007). *English Phraseology: A Coursebook.* Germany: Narr Francke Attempto Verlag GmbH + Co. KG.

Granger, Sylvaine. Paquot, Magali. (2008). "Disentangling the phraseological web," in *Phraseology. An Interdisciplinary Perspective*: 27-50.

Hasselgren, Angela. (1993). *Right words, Wrong words and Different words: an investigation into the lexical coping of Norwegian advanced learners of English*. Unpublished "hovedfag" thesis: University of Bergen.

Hasselgren, Angela. (1994). "Lexical teddy bears and advanced learners: a study into the ways Norwegian students cope with English vocabulary," in *International Journal of Applied Linguistics* 5 (2): 237-260.

Hasselgård, Hilde. Johansson, Stig. Lysvåg, Per. (2007). *English Grammar: Theory and Use*. Oslo: Universitetsforlaget AS.

Henriksen, Birgit. (1999). "Three dimensions of vocabulary development" in *Studies in Second Language Acquisition* 21: 303-317

Jackson, Howard. Amvela, Etienne Zé. (2007). *Words, Meaning and Vocabulary.* London & NY: Continuum

James, Carl. (1998). *Errors in Language Learning and use: Exploring Error Analysis*. NY: Addison Wesley Longman Limited.

Johansson, Stig. (1978). *Studies of Error Gravity: Native reactions to errors produced by Swedish learners of English*. Sweden: Acta Universitatis Gothoburgensis.

Johansson, Stig. (2008). *Contrastive analysis and learner language: A corpus-based approach.* Oslo: University of Oslo

Llach, María Pilar Agustín. (2011). *Lexical Errors and Accuracy in Foreign Language Writing*. UK; USA; Canada: Multilingual Matters

López, Willelmira Catillejos. (Date of publication not specified) "Error Analysis in a Learner Corpus. What Are the Learners' Strategies?" accessed:

http://www.um.es/lacell/aelinco/contenido/pdf/45.pdf 24.04.2013

Martin, Marilyn. (1984). "Advanced Vocabulary Teaching: The Problem of Synonyms" in *The Modern Language Journal*, 68: 130–137.

McEnery, Tony. Hardie, Andrew. (2012). *Corpus Linguistics*. UK: Cambridge University Press.

Meara, Paul. (2009). *Connected words: Word associations and second language vocabulary acquisition*. Amersterdam: John Benjamins Publishing Co.

Ortega, Lourdes. (2009). *Understanding Second Language Acquisition.* UK: Hodder Education.

Pawley, Andrew. Syder, Frances Hodgetts. (1983). "Two puzzles for linguistic theory: native-like selection and native-like fluency" in (Red) Richards, S.C. And Schmidt, R. W. *Language*

and Communication: 191-226.

Rayson, Paul. Log-likelihood calculator website. http://ucrel.lancs.ac.uk/llwizard.html accessed 01.05.2013

Richards, Jack C. (1976). "The role of vocabulary teaching," in *TESOL Quarterly*, vol. 10, no.1: 77-89

Ringbom, Håkan. (1987). *The Role of the First Language in Foreign Language Learning*. UK: Multilingual Matters Ltd.

Saeed, John I. (2009). *Semantics*. UK: Blackwell Publishing.

Schmitt, Norbert. (2000). *Vocabulary in Language Teaching*. UK: Cambridge University Press.

Schmitt, Norbert. (2010). "Key Issues in Teaching and Learning Vocabulary," in Chacón-Beltràn, Rubén; Abello-Contesse Christián; Torreblanca-López María del Mar. *Insights into Non-native Vocabulary Teaching and Learning*: 28-40.

Simensen, Aud Marit. (1998). *Teaching a Foreign Language*. Bergen: Fagbokforlaget Vigmostad & Bjørke AS.

Sinclair, John. (1991). *Corpus Concordance Collocation.* UK: Oxford University Press.

Taylor, Barry. (1975). "The use of overgeneralization and transfer learning strategies by elementary and intermediate students of ESL" in *Language Learning* 25: 73-107

Vasiljevic, Zorana. (2010). "Word Associations as a Vocabulary Teaching Strategy in an Advanced L2 Reading Class," in Chacón-Beltràn, Rubén; Abello-Contesse Christián; Torreblanca-López María del Mar. *Insights into Non-native Vocabulary Teaching and Learning*: 41-61.

Wood, David. (2001). "In Search of Fluency: What Is It and How Can We Teach It?" in *The Canadian Modern Language Review* 57 (4): 573-589.

APPENDIX 1

National winter mock exams in English, 2012/2013. Distributed to 10th graders in Norway

Oslo kommune
Utdanningsetaten

Felles heldagsprøver 2012/13

10. trinn

Oppgave

Fag: Engelsk

Prøvedag: fredag 30. november 2012

In this mock exam there are three tasks.

You must do all three tasks.

Task 1

In Part A you were asked to focus on loving the job, loving ourselves or loving the planet.

Which text in your preferred section is your favourite and why? Write a couple of paragraphs.

Task 2

In Part B you have read two different stories – both of which focus on staying alive:

"Nailer is hanging on to life by a thread. Henrietta, on the other hand, died young, but her cells have given life to thousands of people."

Comment on the quote from *Focus Magazine*. Write a couple of paragraphs.

Task 3

Choose ***ONE*** of the tasks from 3A to 3E. A longer answer is now required.

If you want to, you can use information, ideas, useful words and phrases etc. from the texts you worked with while preparing for this test.

Task 3A

Write a text to fit in with one of the *Focus Magazine* sections:

Focus on Loving the Job
or
Focus on Loving Ourselves
or
Focus on Loving the Planet

- State which section you are writing for.
- Choose and state your genre.
- Give your text a suitable title.
- Refer to at least one person or text from *Focus Magazine.*

Task 3B

In *Focus Magazine* you have read about many different people, real and fictional. Imagine that you are to spend a weekend with one of the characters. Write a text about your weekend.

- Choose and state your genre.
- Create a believable setting.
- Include relevant information from *Focus Magazine*.
- Give your text a suitable title.

Task 3C

In the book *Ship Breaker* Nailer Lopez lives in a future setting with dramatic climate changes because people have been careless about taking care of and respecting Mother Nature.

In most computer games, the choices you make, will affect the outcome. Imagine you have just started to work for a computer game company. They want you to produce a new concept.

Write a proposal where you among other things:

- Explain what the computer game is about and how you can win.
- Explain the consequences of the various options the game offers.
- Explain whom the game is designed for.
- Include at least one person from *Focus Magazine* as your contestants.
- Use the name of the game as your title.

Task 3D

The Ship Breaker, which you read about in part B, tells the story about how climate changes have affected our world and turned it into a world where people are struggling to survive. Write a story set in a future Norway, where climate changes have altered living conditions significantly.

- Create a believable setting in time and place.
- Refer to at least one person or text from *Focus Magazine*.
- Give your text a suitable title.

Task 3E

In both part A and B you have read about people who are passionate about something; their jobs, our planet, their appearances or simply just staying alive. What are you passionate about? Write an article.

- Give reasons for your answer.
- Include relevant information from *Focus Magazine*.
- Give your text a suitable title.

APPENDIX 2

Overview of questions/topics for the NICLE corpus (Johansson 2008: 117)

- Abortion: right or wrong?
- All armies should consist of entirely professional soldiers: there is no value in a system of military service
- A man/woman's financial reward should be commensurate with their contribution to the society they live in
- Crime does not pay
- Europe
- For and against breakfast
- The role of censorship in Western Society
- Feminists have done more harm to the cause of women than good
- Marriage – an anachronism
- In his novel Animal Farm, George Orwell wrote "All men are equal, but some are more equal than others." How true is this today?
- In the 19th century Victor Hugo said: "How sad it is to think that Nature is calling out but humanity refuses to pay heed." Do you think this is still true today?
- Marx once said that religion was the opium of the masses. If he was alive at the end of the 20th century, he would replace religion with television
- The prison system is outdated. No civilized society should punish its criminals: it should rehabilitate them
- Most university degrees are theoretical and do not prepare students for the real world. They are, therefore of very little value.
- Some people say that in our modern world, dominated by science, technology and industrialization, there is no longer a place for Dreaming and Imagination.

APPENDIX 3 (Intermediate Dissonances)

All dissonances found in the **intermediate** student texts

Total: 704 tokens, 301 types **L1 influence**: 434 tokens

Explanation of symbols

Bold – dissonance

"Quotation marks" - (Possible) Norwegian L1 influence

Italics – (Suggested) correct English L2 item (There could be many more alternatives than suggested)

(Parentheses) - Examples from text to illustrate dissonance in context (Examples are largely unabridged, meaning that typos, grammatical mistakes and other features are transcribed faithfully)

[Square brackets] – Words that can also be included in the L1/L1 item, but may not be in all contexts

X - No influence (Norwegian)[13] / No equivalent L2 item (English)

Ø - Item is avoided, either by restructuring clause or using other, more appropriate words

C – Core word

t – Lexical teddy bear

++ - If the dissonance could be replaced by a wide variety of other L2 items

Numbers

Numbers **before** the dissonance tell us how many tokens were found for this one type. If there is no number specified, this means only one token was found.

Numbers **after** the dissonance inform us the sense for the word. If one word has several senses, they will be numbered chronologically: "1, 2, 3" etc.

Ex. 13 **Get** 2 (13 tokens for "get," and this is the second sense listed for "get")

13 Although something is marked as x from Norwegian to English, it does not mean that the structure does not exist in Norwegian, simply that it was not an influencing factor in why the speaker chose the wrong word

COGNATES

97 tokens, 45 types L1 influence: 97 tokens

Cognate/Semantic

42 tokens, 13 types L1 influence: 42 tokens

Consistence - "konsistens" - *consistency*
(and he could feel the sticky **consistence** of sweat all over his body)
3 **Clock -** "klokken" - *Time/it*
(He had to hurry up, the **clock** was almost eleven)
2 A **cut -** "ett kutt" - *gash*
(A man who got a big **cut** on his left arm)
2 History – historie – s*tory*
(I think that both of the **histories** tell us a very good story)
Hold – "holde" - *keep*
(He **held** his rifle down)
Job- "jobb" - *process/act*
(The **job** of making your body perfect)
2 A **land** - "et land" - *country*
(Almost no other **lands** were trading with us)
To **mean that -** "å mene" - *to believe/be in the opinion of/think that*
(I **mean** that both of them try to save themself from something evil)
Meet - "møte" - *face*
(This issue is something we **meet** in our everyday life)
Placed - "Plassert" - *Located*
(The town was Yoleville, **placed** in the very west of New Jersey)
8 To **see** - "se" - *look*
(She **saw** me in the eyes/ I **saw** myself in the mirror)
4 To **train**- "trene" - *work out/exercise*
(Some of them **train** so they loose allot of weight)
2 **Way** - "vei" - *path/route*
(Others choose another **way**)

Cognate/Collocational

22 tokens, 10 types L1 influence: 22 tokens

A **baby** voice - "en baby stemme" - *childish* voice
To **come** anywhere - "komme noensted/vei" - *go* anywhere
Give an opinion - "gi en mening" - *offer* an opinion
A **glad** voice - "en glad stemme" - *cheerful/happy* voice
2 To **have** right/the impression - "ha rett/inntrykket" - *be* right/*be under* the impression
Little summary **-** "liten oppsummering" - *brief* summary
2 To say something **low** - "å si noe lavt" - *softly/quietly*
11 **Take** surgery/vaccine – ta kirugi/vaksine – *have* surgery/vaccine
To **take** on clothes - "ta på klær" - *put* on clothes
See what **time** brings - "se hva tiden vil bringe/vise" - see what *the future* will bring

Cognate/Stylistic/Connotational

17 tokens, 8 types L1 influence: 17 tokens

Connotational

2 **Cold -** "kald" - *cool*
(Luckily the **cold** air-condition was on)
Fantasy character - "fantasifigur" - *fictional* character

Informal

2 **Come up with** – "komme opp med" – *think of*
(Someone will **come up with** something)
Find – "finne"– *discover*
(No one has **found** oil)
4 **Fix** – å fikse [på] – *alter/change*
(**Fix** their appearance)
3 **Much** – "mye" – quite a bit
(So **much**/very **much**)
3 Say – si – s*tate*
(As **said** in the quote/the magazine **says**)
So-called – "såkalt" – Ø
(And a gruesome sound, or the **so called** ''biip'' went over the whole room)

Cognate/Syntactic

5 tokens, 4 types L1 influence: 5 tokens

To describe very **good** – veldig godt – very *well*
The **medicine** world – medisinverdenen – *medical* world/*medicinal* world
2 A **sew** job – en syjobb – a s*ewing* job
Very much rain – veldig mye regn – *a lot of* rain

Cognate/Invalid

11 tokens, 10 types L1 influence: 11 tokens

To **come in on** - "komme inn på" - x (*get* into/*are accepted* into)
(You **come in on** nurse school)
Day to day life person - "dagliglivsperson" - x (everyday/common person)
(That isn't the case for a normal **day to day life** person)
A **drastic** attempt – "et drastisk forsøk" - x
To **give a wink** - "gi et vink" - *wave*
2 To **hold on** to his/her life - "holde på livet" - *stay alive*
A **prettier** body - "penere kropp" - x
Psychical suffering - "psykiske lidelser" - *mental illness/suffering*
A **sour** wind - "sur vind" - x
Take [someone's] **life -** "[noe] tar livet hans/hennes" - x (something/one kills a person)
(This [accident] almost **took** his **life**)
Take oneself through - "å ta seg gjennom" - *go* through/*put* oneself through
(He is not thinking of the danger that he is often **taking** him through)

TRANSLITERATIONS

93 tokens, 68 types L1 influence: 93 tokens

Transliteration/Semantic

15 tokens, 12 types L1 influence: 15 tokens

Allow - "å la" - *let*
(She **allowed** her skinny legs to carry her to the bathroom)
The theme **appears** because … -"dukker opp" - is present
Case – "sak" - issue/subject
(I like to give my opinions […] and let people know how I feel about the **case**)
To die **early** - "dø tidlig" - die *young*
(It then mentions that Henrietta *on the other hand*, **died early**)
You can only lose life on the **different** side – "andre siden" - *other* side
Funny - "morsom" - *fun*
(after a **funny** day)
3 To **get away** - "få vekk" - *remove*
(Use botox to **get away** rankles in their face)
To grow – å vokse – gained popularity/become more common
(Plastic surgery has **grown**)
2 To **lay [**something] down - "legge noe ned" - *put* down
To **live** - "å bo" - *stay*
(We had first planed to **live** at a hotel)
To be **locked** inside - "låst inne" - *trapped* inside
Look at – "se på" - *perceive/view/regard*
(They **look at** it as an attack)

Transliteration/Collocation

22 tokens, 18 types L1 influence: 22 tokens

A plane **accident -** "flyulykke" - plane *crash*
Boring sky - "kjedelig himmel" - *dull* sky
To **close** the door - "lukke" - *shut the door*
Life **contains** -"livet inneholder" - life as a [something] *means*
(I learned more about what the life as a nurse can **contain**)
2 A **dark** voice - "dyp stemme" - *deep* voice
The world has **developed** into - "utviklet seg" - *become* (?) / *turned into*
To **do** a good **try** - "gjøre et godt forsøk" - *make a valid attempt*
To **drive** a cab - "å kjøre taxi" - *take* a cab
An **evil** circle - "en ond sirkel" - a *vicious* circle
Take a **fast** look - "ta en fort titt" - Take a *quick* look
Inflict damage - "påføre smerte" - *cause* damage
She **held** her ears - "holdt for ørene" *clamp* her ears
4 To **make** - "lage" - *cook* food/*build* a plane/*throw* a birthday party
To **order** a **plane** - "bestille et fly" - *book a flight*
Positive **parts** - "positive deler" - positive *sides*

Showed up in her mind - "dukket opp i tankene" - *appeared* in her mind
A car **standing** in the driveway - "en bil står" - *parked* in the driveway
Throw light on - "kaste lys på" - *shed* light on

Transliteration/Stylistic/Connotational
43 tokens, 26 types L1 influence: 43 tokens

Connotational
3 **Failed** surgery - "mislykket operasjon" - *botched* surgery
Large people - "Store mennesker" - *overweight/obese* people
(Many teen age girls think that the fact that you are perfect like you are, is some Pease of skit that **large** people says to them selves)

Informal
7 **Actually** - "faktisk" – *in fact*/Ø
(**Actually** too big)
All kinds of – "allslags" – *Ø/all types*
(Remove **all kinds of** small covers and such)
Always – "alltid" – Ø
(I have **always** loved fantasy)
Anyway – "uansett" – after all/Ø
(It was too late **anyway**)
Play a **big** role - "spille en stor rolle" - play a *major* role
2 **Funny** – "morsom" – *humorous/entertaining*
(It was very **funny**)
2 **Get** 8 – "få" – *are permitted/allowed*
(We **get to** learn)
2 **Get the feeling** – får følelsen– Ø
(You **get the feeling** that the Queen wants to be perfect)
Give an answer – "gi et svar" – *reply*
(You can **give** me **an answer**)
Go as planned – "gå som planlagt" – *Go according to plan*
(Some of the teenagers begins to cry if not birthday **go as planned**)
Go so well with – "gå så bra med" – *is successful ++*
(You can find thousands of pictures of people that it hastens **go so well with**)
Go wrong – "gå galt" – *Ø/fail/be unsuccessful*
(Can very easily **go wrong**)
A mistake **happened** - "en feiltagelse hendte" - a mistake *occurred*
Important for saving lives – viktig for å redde liv – *vital* for saving lives*
It hit [someone] – "det slo [noen]" – *make someone realize*
(But then it **hit** me, even though)
2 **Pretty** 1 (adverb) – "ganske" – *quite*
(It was **pretty** good)
Something like that – "noe sånt" – Ø
(Never do **something like that** agene)

Something to do with – "noe å gjøre med" – *Ø/in common/are related to*
(Both here have **something to do with** saving)
Talk about– "snakke om" – *discuss*
(It`s this I`m going to **talk about** now)
That's right – "det stemmer" – Ø
(Well, **that's right**)
4 The text **told** - "teksten fortalte" - the text *informed*
2 **Totally** – "helt" – completely/ vastly different
(**Totally** different method)
2 **Way** – "måte" – *manner**
(There are many different ways to show love in the **same way** as it's easier to)
Weird – "rart" – strange*/curious/odd/peculiar
(It looked **weird**)
Transliteration/Syntactic
3 tokens, 2 types L1 influence: 3 tokens
Misinterpreted **easy** – lett – *easily*
2 Look [tired/angry] at [someone] - "se [trøtt/sint] på noen" - look at someone tiredly/angrily
(He looked **tired** at them)

Transliteration/Invalid
10 tokens, 10 types L1 influence: 10 tokens
Building people - "byggefolk" - *construction workers*
Broken into pieces - "knust i biter" - *torn to bits*
An **eating** table - "et spisebord" - a *dining* table
An eating **disturbance** - "spiseforstyrrelse" - *eating disorder*
To look **empty** out the window - "stirre tomt ut vinduet" - x
Time **ran** from her - "tiden løp fra henne" - x
A **safe** bed - "en trygg seng" - x
Shortly said – "kort sagt" - *briefly put*
A **strict** look - "strengt blikk" - x
A **thin** body shape - "tynn kroppsform" - x

PERCEIVED EQUIVALENTS
21 tokens, 14 types L1 influence: 0 tokens

Perceived Equivalent/Semantic
19 tokens, 12 types L1 influence: 0 tokens
A **better curve** – x – unknown intended meaning
(Most of the girls that take plastic surgery are mostly focusing on getting bigger lips, **a better curve** or a facelift)
Give me an answer if we want to make this a **big deal** – x – unknown, possibly "make an agreement"
3 To **catch** - x- to *understand*
(She would **catch** it like that)
3 To **evolve** - x - to *develop*

(We didn't have the money to **evolve** something to get rid of the volcanoes)
To **expose** – x - *Unknown intended meaning*
(A moment of silence **exposed** itself)
A **mortal** glance – *Unknown intended meaning*
Make someone **noble** – x - *Unknown intended meaning*
Payback – x – *Unknown intended meaning*
(And get **payback** for all the training has did to get into that size)
2 A **Profit** - x– *Goal/intention*
(Maybe the **profit** is to reach popularity)
2 **Regular -** x - *usual/common*
To v**isualize** – x – *Unknown intended meaning*
(But also a lot of other things you have to **visualize** before you are going out on a trip)
2 To give a **withering** look – *x – Unknown intended meaning*

Perceived Equivalent/Invalid
2 tokens, 2 types L1 influence: 0 tokens
Job community – x -*Work force*
(So to be happy in the **job community** you go for, is)
She said **monotony** to herself

SYNONYMS
130 tokens, 102 types L1 influence: 13 tokens

Synonym/Semantic
24 tokens, 20 types L1 influence: 1 token
I have been remembering **all** day since – x – *every* day
Apparently innocent blue eyes – x – *seemingly* innocent
Backside – x – *downside*
(There is a **backside** though)
More than adults **can** ever know – x – *could* ever know
We went down to the **coast** to get on his boat – x – the *shore*
The **divisions** of a hospital - "avdeling" - *wards*
They **don't** take that much damage –x – *can't* handle
I **don't** tolerate this – x – I *won't* tolerate
The **environment** around him stood still - x - *surroundings*
I **grasped** my pen out of my pocket – x– *grabbed/took out/pulled*
It is not **immoral** to do some changes or improvements – x– *bad/negative*
To **name** – x – To *call*
(and that includes fixing our body, also **named** as plastic surgery)
I asked if she had time to **pick** me **up** to work –x– *drive me to work/hitch me a ride*
He **raised up** some binoculars from his bag – x – *grabbed/took/pulled out/fetched*
Those lives couldn't **stand up** against the one he had ruined – x – *amount to*
Translated to - x - *interpreted* as
(This can be **translated to** the thread he lives on)
When the shipment was ready, he **warned** other lab workers – x – he *notified*

5 A **wood** – x – a *forest*
(Her uncle Brian owns a **wood**)
I **would** give you the answer now – x – I *will* give you

Synonym/Collocational
35 tokens, 33 types L1 influence: 9 tokens
When the plane **arrives** - "når flyet ankommer" - plane *lands*
The **bottom** theme – x – the *underlying* theme
A **broken** ship – x – a *wrecked* ship
Taking no **care** – x – taking no *heed*
Choose the right decision – x – *make* the right decision
A car **coming** closer – x – *pull up* closer
How **could** I know? - "hvordan kunne jeg vite det?" - How *should* I know
I worked to **delete** it from the earth - "slette noe fra verden" - *erase/eradicate* it
The volcano **detonated** – x – the volcano *erupted*
Sleeping **drugs** – x – sleeping *pills*/*sedatives*
On the **edge** to die – x – on the *verge* of dying
Everyone **escaped** into the mountains - "rømte til fjells" -*fled* to the mountains
Lose a **fight** - "tape kampen" - lose a *battle*
A job that **fits** you - "en jobb som passer deg" - *suits* you
Go active – x – *become* active
(Volcanoes **went** active)
Going on a mountain – x – *walking* on mountains
Go out of bed – x – **get** out of bed
(I **went** out of bed)
Good and downsides – x – *upsides* and downsides
A **happy** job – x – A *good/satisfactory* job
A smile crossed her **mouth** – x – crossed her *lips*
Good thing/**negative** thing – x – good/*bad*
In **nice** shape – x – in *good* shape
The **nice** side – x – the *bright* side
I remember it like it was **the previous** week – "sist uke" – *last* week
2 A **quick** move - "en rask bevegelse" - A *swift* move
To **quiz** a survey – x – to *take* a survey
To **reach** popularity – x – *gain* popularity
To **squeeze** a needle – x – to *inject* a needle
Surrounding that question – x – *concerning* that question
2 Take the warning – x – *listen to/heed* the warning
Tell a speech – x – *hold* a speech
The water almost **touched** the roof tops - x - *reached* the roof tops
Use a plane – x – *take* a plane

Synonym/Stylistic/Connotational
62 tokens, 43 types L1 influence: 2 tokens
<u>Connotational</u>

She was too busy [eating] the **cadaver** – x– too busy [eating] the *dead body*
A **regular** flood – x – an *ordinary* flood
The magazine should be more **correct** – x – *accurate*
Slightly horrendous – "litt fælt" – *quite/terribly/incredibly*

Formal
Execute your job the best way possible – x – *perform/do* your job
Important for a **human** to have a job s/he enjoys – x – a *person* to have a job
Everyone ran to the **human being** – x – the *person*
Instead of things that are **sinister** and **dismal** – x x – *negative/bad* and *depressing/sad*

Informal
Another thing – x – Ø
(**Another thing** I think is very good)
A ton – x – Ø
(Better with no make up, than **a ton** of it)
2 **Basically** – x – Ø
(**Basically,** people take plastic surgery)
4 **Boobs** – x - *breasts*
(**Boob job**/ big **boobs**)
2 **Boring** – x – *dull/mundane/meaningless*
(The world would be a **boring** place to live)
Fake abs – x - *abdominal etching*
2 **Fat** – x – *overweight/obese*
(They still look **fat**)
2 **Figure out** – x – *understand/comprehend/Ø*
(I couldn't **figure out** what they had to do with focus on staying alive)
Finish up – x – *complete*
(I had to **finish up** the last five months of junior high)
2 **Get** 5 – x – *go/travel*
(To **get** from one place to another)
2 **Get** 9 – x – *enter/arrive*
(Their main concern was to **get to** heaven)
Get more into – x – Ø
(You will **get more into** what he think about his life)
2 **Get to know** – x – *to become acquainted with/learn*
(**Getting to know** people at the same time as executing your job)
Go downhill– x – *Ø/go poorly/not well*
(The Norwegian economy had **gone downhill**)
2 **Go for** – x - *pursue*
(When the people that don't know what job they should **go for**)
2 **Gross** – x – *unpleasant*
(Working with injured people might seem **gross**)
A **huge** lump in my throat - "en stor klump i halsen" - a *large* lump
Get **hurt** by a knife – x – get *injured* by
2 **Keep [on]** – x – *continue*

(Her cells **keep** living after her tragically early death)
Kind of 1– x – *type of/sort of*
(It was a good **kind of** chocolate)
Let it go – x – Ø
(I guessed it was a routine question so I **let it go**)
Literally – x – Ø
(They tried the put some cells in people who were hurt in accidents, and needed a piece of meat, **literally**)
2 **Or something** – x – Ø
(While she watched Animal Planet **or something**)
Piece of shit – x – Ø
(Is some **pease of skit** that large people says to them selves)
Put it– x – *express*
(That is exactly how I would have **put it**)
3 **Right** – x – *correct*
(This statement is **right**)
Stick to – x – *stay/perservere*
(**Stick to** that job for the rest of your life)
Stuff – x – Ø
(It was **stuff** like)
Stupid – x - *foolish*
(Her **stupid** dad)
The good news – x – Ø
(Midge goes back to the wood to tell the little people **the good news**)
The whole world – x – Ø
(Besides being the most beautiful girl i **the hole world**)
Girls my age is **upset** because of their body –x – *discontent*
4 **Way** [too many/few/above] – x – Ø
(**Way too** many plastic surgeries / **Way** above the volcano)
You see – x – Ø
(**You see**, we are going to work)

Synonym/Syntactic
7 tokens, 4 types L1 influence: 0 tokens
An airplane accident my friend **drove** – x – x
4 **get** born/die – x – *are* born/Ø die
She **fetched** a deep sigh – x - x
They like each other very **well** – x – very *much*

Synonym/Invalid
2 tokens, 2 types L1 influence: 1 token
New mown grass - "nyklippet gress" - *fresh(ly) cut* grass
Burning **damage** - x - burn

ASSOCIATIONS

52 tokens, 47 types L1 influence: 1 token

Association/Semantic

16 tokens, 14 types L1 influence: 0 tokens

Not **also** –x - *not only*
(Not **also** have they been very unlucky with the surgeries)
To **approach** – x – to *approve*
(Yes, you can do that. I would **approach** that)
To **arrive** – x – to *appear*
(A new planet had just **arrived** very close to the Earth)
Defiantly – *x* – *definitely*
Two **dissimilar** ways - x - two *different* ways
Effect – x - *affect*
(It **effects** the rest of your life)
Existing – x – *excitement*
Front head – x – *forehead*
Her arms were **latent** on each side of her – x – lying
Living – x – *leaving*
(I say "yepp lets go... (shouting) by mom I am **living**")
Prepositions – x – *proportions*
(Or you got born with the **prepositions** in your face wrong)
Raise – x – *rise*
(The ocean **raised**)
3 To **select** money – x - to *collect* money
Taught – x – *thought*

Associational/Collocational

9 tokens, 8 types L1 influence: 1 token

2 When worlds **crash** – x – worlds *collide*
Consisting picture – x – *containing* pictures
An **exaggerated** birthday party – x – *extravagant* birthday party
To **increase** your look – x – to *enhance* your look
Reduce beauty pressure – x – *relieve* beauty pressure
To **search** for approval - "lete/jakte etter godkjennelse" - *seek* approval
She let her hand **rush** through her hair – x – *run* through her hair
The magazine **posted** – x – *published*

Association/Syntactical

21 tokens, 20 types L1 influence: 0 tokens

Addict – x - *addictive*
An **argue** – x – an *argument*
Her hair looked **beautifully** – x – *beautiful*

Can food – x – *canned* food
It was **completely chaos** – x – *complete* chaos/*completely* chaotic
To live a bit **complex of** day to day life – x – to live a *complex/complicated* life
Love is **difference** – x - *different*
She was very **dramatically** – x - *dramatic*
Near **dying** – x – near-*death*
Surgery **fails** – x – surgery *failures*
A **fiction** character – x – *fictitious*
She said **flirty** – x - *flirtatiously*
It was not **necessarily** – x – *necessary*
Feeling of **perfect** – x – feeling of *perfection/being perfect*
A **reality** character – x – *real/based on real life*
Seriously – x – *in all seriousness*
2 To hold a **speak** – x – a *speech*
Fight for **survive** – x – *survival*
Technology upgrades – x – *technological* upgrades
A **teleport** spell – x – a *teleportation* spell
Association/Invalid
6 tokens, 5 types L1 influence: 0 tokens

To get **beater** – x – *get better*
2 **Rankles** – x – *wrinkles*
He **sward** – *x – he swore*
To **tare** through flesh – x - *tear*
Wreaked – x – *wrecked*

CORES AND LEXICAL TEDDY BEARS
311 tokens, 25 types L1 influence: 230 tokens

Core/Semantic
1 token, 1 type L1 influence: 0 tokens
They **have got to** explain very well – x – they *explained*

Core/Stylistic
310 tokens, 24 types L1 influence: 230 tokens

7 **A little bit [of]** – "litt" – *a little/slightly/some/Ø* t
(**A little bit** entertaining)
20 **A lot [of]**– "mye" – *quite a bit/much* t
(Do it with **a lot of** care)
2 Bad – x – *unfortunate* **C**
(The only **bad** thing)
28 **Big** – "stor"– *large* **C**/t
(Too **big**)
7 **Deal with**– x – *cope with* t

(It is difficult to **deal with** situations like death alone)
5 **Do [anything/everything/something]** – "gjøre hva some helst/alt/noe" - Ø/sacrifice **t**
(She is ready to **do anything** to get prettier)
23 **Get** 1 – x – become/be/Ø **C/t**
(**Get** thinner/ my vision **got** blurry/The word **got** smaller)
22 **Get** 2 – "få" – *receive/Ø/find/obtain* **C/t**
(Where they **get** mental help/All they could **get**)
7 **Get** 3 – "få" – *have* **C/t**
(You **got** cancer)
20 **Good** – "god"– *wonderful/positive* ++ **C/t**
(How **good** it was)
4 **Great** – x – wonderful/excellent/ ++ **C**
(**Great** body, **great** time)
5 **Huge** – "stor" – *large/overwhelming* ++**t**
(A **huge** responsibility)
8 **Kind of 2** – "på en måte" – *in a way/can be perceived as/Ø* **t**
(**Kind of** ironic)
8 **Like** – "like" – *enjoy/appreciate/Ø* ++ **C/t**
(I **like** the way)
6 **Look** 1 (verb) – "se ut" – *appear* **C/t**
(To help them **look** better/everything looks so easy)
12 **Look** 2 (noun) – "utseende" – *appearance** **C/t**
(A pretty **look**)
7 **Make** – "lage/få" – *create/suit better/ causes/Ø* **C/t**
(This was **made** for girls /The virus **makes** humans die)
4 **Nice**– x – wonderful/lovely ++ **C**
(She looked so **nice**)
17 **Pretty 2** (noun)– "pen" – *attractive/beautiful* **C/t**
(Girls worry about being **pretty**)
13 **Really** – "virkelig" – quite, truly **C/t**
(I'm **really** glad)
7 **Sad** – "trist" – *tragic/unfortunate* **C/t**
(A **sad** story)
20 **Thing** – x – Ø **C/t**
(The **thing** about)
20 **Think** – x – *find/in the opinion of* **C/t**
(I **thought** the text was good)
38 **Very** – "veldig" – quite/extremely ++ **C/t**
(Very interesting)

APPENDIX 4 (Advanced Dissonances)

All dissonances found in the **advanced** student texts

Total: 373 tokens, 212 types **L1 influence**: 200 tokens

COGNATES
48 tokens, 25 types L1 influence: 48 tokens

Cognate/Semantic
12 tokens, 8 types L1 influence: 12 tokens
Break human rights - "bryte menneskerettigheter" - *violate* human rights
To **can** - "å kunne" - *know*
(A chance to show what they **can**)
Learn - "å lære" - to *teach*
(He is not able to **learn** the children what he knows)
To **group** with - "å gruppere (sammen) med" - *associate*
(We **group** with those we feel we have most in common with)
4 To **mean** – "å mene [at]" - to *believe/think/be in the opinion of*
(Others **mean** that a person with a physical or mental defect should be spared from living …)
Psychic – psykisk – *psychological/mental*
(When it comes to physics and and/or on the **psychic** level)
2 **Shall** – "skal" - *should*
(Why **shall** some people have more money than others?)
Way - "vei" - *path*
(Somehow the **way** from education to work seems)

Cognate/Collocational
12 tokens, 6 types L1 influence: 12 tokens
Give a punishment – "å gi en straff" - *hand out* a punishment
(The punishment **handed out** by the authorities)
3 **Live** a life – "leve et liv" – *lead* a life
(Improving the life we actually do **live**)
Much or **less** – "mye eller lite" - *a lot* or *a little*
(It is your own responsibility whether you earn **much** or **less**)
2 On the other **side** – "på den annen side" - on the other *hand*
Strong resources - "sterke ressurser" - *rich* resources
(Those with the **strongest** resources)
4 **Take** an abortion/decision/job - "ta" - *have* an abortion/*make* a decision/*accept* a job

Cognate/Stylistic/Connotational
18 tokens, 5 types L1 influence: 18 tokens
4 **Say** –"si" – s*tate*
(Marx once **said** that religion was the opium of the masses)
8 **Fantasy/fantasize** – "fantasi/fantasere" – *imagination/imagine/think about*
(It can be relieving to **fantasize** and dream about fictive stories)
2 **So-called** - "såkalt" - Ø
(In 1787, the **so-called** Founding Fathers wrote the American constitution)
2 **Sick** - "syk" - *mentally ill*
(An unstable and **sick** person)
2 A **thinking** robot – "en tenkende robot" - *sentient* robot
(The **thinking** robot was once a dream)

Cognate/Syntactic
5 tokens, 5 types L1 influence: 5 tokens
To **argument** – "å argumentere" – *argue*
(Some even **argument** that)
Discriminating – "diskriminerende" – *discriminatory*
(They are just as **discriminating** as they accuse the men for being)
Man dominated – "mannsdominert" – *male*-dominated
(So in a **man** dominated business world)
Much – "mye" – *a lot/plenty*
(People have **much** because of modern scientific instruments)
A **thinking**-process – "en tenkeprosess" – a *thought* process

Cognate/Invalid
1 token, 1 type L1 influence: 1 token
Psychically – "psykisk" – *mentally*
(Will have more practice both physically and **psychically**)

TRANSLITERATIONS
74 tokens, 53 types L1 influence: 74 tokens

Transliteration/Semantic
8 tokens, 3 types L1 influence: 8 tokens
A sign that everything is in order – "et tegn på at alt er i orden" - x (everything is alright)
(Everything is straightened up and you have **a sign that everything is in order**)
Contains – "inneholder" – *consists of*
(A totally different lifestyle, which **contains** fame, publicity and luxury)
6 Look upon/at - "se på" - *regard/view/consider*
(All nations **look upon** each other as independent/ Sometimes they are **looked at** as the outsiders)

Transliteration/Collocation

15 tokens, 12 types L1 influence: 15 tokens

Be a part – "være en del av" – *play* a part

(Money **is** an important part in this game)

2 To **become** eighteen – "å bli atten" - to *turn* eighteen

(When you **have become** 18-19 years old, you still do not know …)

Big demand - "stort krav" - *high* demand

(A society that sets **big** demands)

Cut weeds – "kutte vekk ugress" - *hack* weeds

(You do not want to see the weeds, so you **cut** them off)

Develop an **idea** – "å utvikle en ide" – to *adopt* an idea/develop an *attitude*

(Our modern society has **developed** a very wrong idea)

2 Practical **exercises** - "praktiske oppgaver" - practical *tasks*

(The universities had practical **exercises** with grades)

Know as your own back pocket – "kjenne som sin egen lomme/lommekjent" - *know like the back of your hand*

(The home guard, for instance, **knows** their district **as their own back pocket**)

As **late** as yesterday - "så sent som i går" - as *recent* as yesterday

(I have watched a great deal of television and as **late** as yesterday I watched an episode)

Make a revolt - "gjøre opprør" - *Ø revolt*

(As a result, they did not **make** any revolts against the capitalists)

Use or **misuse** – bruke eller misbruke – use or *abuse*

(Maybe the child will be used or **misused**)

2 **Receive** access/knowledge - "få tilgang/kunnskap" - *gain* access/knowledge

(To prevent non-European immigrants to **receive** access to the country)

White-skinned people - "hvithudet mennesker" - *light* skinned people/*Caucasians*

(**White-skinned people** tend to fraternize and look down upon colored people)

Transliteration/Stylistic/Connotational

45 tokens, 32 types L1 influence: 45 tokens

4 **Actually** - "faktisk" – *in fact*

(**Actually** too big)

A little bit– "litt" – a little/slightly/some/Ø

(**A little bit** entertaining)

All kinds of – "allslags" – Ø/all types

(Remove **all kinds of** small covers and such)

At least – "i det minste" - Ø

(**At least** we put a lot of effort in)

3 **Bad** – "dårlig"- *poor(ly)**

(No matter how **bad** the conditions you lived under were/ Treat someone **bad**)

2 **By the way** – "forresten" – Ø

(And **by the way** my room looked like a)

Do anything – "gjøre hva some helst" - Ø

(She is ready to **do anything** to get prettier)

4 **Do a [good/bad] job** – "gjøre en [god/dårlig] jobb" -*perform well* ++

(But can all of them **do a good job**?)
Easy way out – "lett utvei" – *simple solution*
(This makes abortion an **easy way out**)
2 [Bad/good] **enough** – "[god/dårlig] nok"– *[in]sufficient*
(But it is not **good enough**)
Find – "finne" – *discover*
(No one has **found** oil)
3 Get a benefit – "få en fordel" **-** *gain* a benefit
Get 4 – "få" – *develop*
(There are many people who **get** bad nerves)
Get 8 – "få" – *are permitted/allowed*
(We **get to** learn)
Get back – "få tilbake" – *have returned*
(This part will be really to hard to **get back** if you should ever regret your actions)
Get [someone's] head filled – får hodet fylt – Ø
(Housewives […] **got their head and ears filled up** with speeches)
2 **Get [something/the most] out of** – "få noe utav" – *receive benefits*
(Then it is up to others to […] and try to **get something out of it**)
Get [someone's] way – "få viljen" – *something goes according to their wishes/Ø*
(If they **got their way**, they would send Norway […] back into the patriarchal stone age)
2 **Kind of**– "slags" – *type of/sort of*
(No mather what **kind of** a prison system)
3 **Look** – "se" – *appear/seems*
(To help them **look** better/everything **looks** so easy)
Mess– rot – *disorder/chaos*
(The economic **mess** is in order)
Obviously – "tydeligvis" – Ø
(**Obviously**, they were wrong)
2 **So many [things]** – "så mange [ting]" – **several/many [things]**
(There are **so many things** you can say about marriage)
Talk about – "snakke om" – *discuss*
(It`s this I`m going to **talk about** now)
3 **Take a look at** – "ta en titt på" – *look at/examine*
(I am going to **take a look at** the unequal financial rewards in society)
Take care of – "ta vare på" – *support*
(There are also people who just can't **take care of** their children)
Take [someone] out – "ta noen med ut" *-bring someone for*
(They should **take her out** for a couple of drinks)
Take over - "ta over" **-** *occupied/gained control over*
(Science technology has **taken over** large domains)
Used to – "vant med" – <u>accustomed to</u>
(They have to be given necessary guidelines away from what they are **used** to)
Want nothing to do with – "vil ikke ha noe å gjøre med" – *does not want to affiliate with*
("One Nation" **wants nothing to do with** Asians)

Transliteration/Syntactic

3 tokens, 3 types L1 influence: 3 tokens

Away from – "vekk/bort fra" - Ø

(They have to be given the necessary guidelines **away from** what they are used to)

The danger is small – "faren er lite for at" – *There is a small risk*

(Even though t**he danger is small** that it will ever happen)

Small part - "liten del" - small part *of the population/fraction*

(It is obvious that only a **small part** can afford this lifestyle)

Transliteration/Invalid

3 tokens, 3 types L1 influence: 3 tokens

Life-lie - "livsløgn" (illusions that a person may build their life around, first coined by Henrik Ibsen)- x

(An average person does need the **life-lie** to function)

Office rat - "kontorrotte" (slang for a person who spend too much time in the office) – x

(Maybe dreams still exist with overworked **office rats**)

Self-decided abortion - "selvbestemt abort" - *abortion on demand*

PERCEIVED EQUIVALENTS

8 tokens, 8 types L1 influence: 0 tokens

Perceived Equivalent/Semantic

8 tokens, 8 types L1 influence: 0 tokens

In **conformity** with – x – unknown intended meaning

(In **conformity** with drugs, smoke and alcohol, television is highly addictive)

Division – x – unknown intended meaning, possibly *gap* in time

(A *division* in time will best justify this matter)

Everything falls to pieces – x – unknown intended meaning, possibly *falls into place*

(Now I have stated the facta about anachronism, **every thing fall to pieces**)

Earlier – x – unknown intended meaning, possibly *former*

As the **earlier** classes in society are more or less evened out)

Existence – x – unknown intended meaning

([Military systems] do resemble each other in more than one field of strategy and **existence**)

Programme form – x – unknown intended meaning, possibly *television genre*

(**Programme forms** such as situational comedies)

(They do resemble each other in more than one field of strategy and **existence**)

Significant – x – unknown intended meaning, possibly *independent/original*

(...Ukraine and Pakistan all have their own **significant** [military] systems)

Stick out – x – unknown intended meaning, possibly *strike*

(He also said that political conditions are transformed in the cultural sphere, so that they do not **stick out** as improper)

SYNONYMS
112 tokens, 91 types L1 influence: 27 tokens

Synonym/Semantic
13 tokens, 11 types L1 influence: 3 tokens
According to – x – *after/as a result of* (?)
(**According to** recent in events in Norway)
Come up – x – *appear/are produced*
(I am truly amazed by all the fancy products from fancy industries that **come up** nowadays)
Divisions – x – *boundaries*
(Which spoke to people across class, ethnic, and gender **divisions**)
Earthly – "verdslig" – *worldly*
(Both religious and **earthly** constrictions)
Essential cause – x – *main/primary* cause
(Greed and excess consumption have often been **essential** causes [for exploiting nature])
Force – "styrke" **-** *strength*
(In the extreme case of war, people will fight with an unknown **force** far greater than …)
Guard – beskytte – *protect*
(Censorhip **guards** leaders against unwanted criticism)
Harmonize – x – *equal*
(They often get a punishment that does not **harmonize** with the crime they have committed)
Matter – x – *factor*
(Age is also a **matter** which creates inequality)
3 **Notorious** (criminal)– x – *minor offender*
(A **notorious criminal** has committed several minor crimes and get eventually caught)
Think – x – *find/in the opinion of*
(I **thought** the text was good)

Synonym/Collocational
25 tokens, 20 types L1 influence: 7 tokens
2 Achieve knowledge – x – *gain* knowledge
(The university as a place to **achieve** their knowledge)
One other **aspect** of seeing it – x – *way* of seeing it
A **better** level – x – a *higher* level
(A **better** level of expertise)
Keyboard **button** – x – keyboard *key*
(Push a few **buttons** on the keyboard)
Call wrath – x – *incur* wrath
(They were afraid they might **call** God's wrath upon them)
Connect bonds – knytte bånd (possible associative influence)– *forming* bonds
(Quality-time spent **connecting** bonds to your children)
2 Do a crime – x – *commit* a crime
3 Female – "kvinne" – *woman*
(The ideal family **female** have been killed)
2 **Fight** – "kamp" – *struggle**
(Little has changed since the **fight** [for women's rights] originated in the 1800s)

A **great** step – x – *large* step
(Would be a **great** step to take)
Happy and **conflicting** times – x – **good** and **bad** times (counts as 2 dissonances)
Human beings - "mennesker" - *mankind*
(For the good of **human beings**)
Human life - x - *life*
(Religion has always been an important part of **human** life)
Hunted – x– *persecuted*
(Those with such views [non-religious] were **hunted**)
Keep to the facts – x – *stick* to the facts
To **pinpoint** a point – x – to *make* a point/to *pinpoint something*
Put someone straight – x – *set* someone straight
(There is no doubt that the military has **put** several people straight)
Put upon – x – *imposed on*
(They would care less about the exploitation **put upon** them by the capitalists)
To take the wrong signal – x – take *something* the wrong *way*
(When a person is not rewarded he or she may **take the wrong signal**)
A **way** of punishment – x – a *form* of punishment

Synonym/Stylistic/Connotational
68 tokens, 54 types L1 influence: 17 tokens
Stylistically too formal
2 **Human beings** – "mennesker" – *people*
(They need to buy new things to be successful *human beings*)
Lady – "kvinne" – *woman*
(In Muslim countries, **ladies** are subdued by men)

Stylistically too informal
A bunch of – x – Ø
(A standing army of professionals can be called out to missions much faster than **a bunch of** nineteen-year-olds)
All over again – x – Ø
(The army lets people start **all over again**)
A long time ago – x – Ø
(This happened **a** very **long time ago**)
A whole lot more – x- Ø
(Modern mobile phones have everything you need, and **a whole lot more**)
Bad – x – unfortunate (this case is also stylistically unacceptable in Norwegian, so is not transliteration)
(The only **bad** thing)
Bad enough as it is – x – Ø
(This world is a **bad enough place as it is**)
2 **Basically** – x – Ø
(**Basically,** people take plastic surgery)
Broke – x – *poor/poverty-stricken*
(To prevent that they are broke when they are "on their own" again)
2 Brainwash – "hjernevaske" – *indoctrinate*

(Tend to be more **brainwashed** and focused on his job)
2 Called up into the army - "innkalt til militæret" - *drafted* in the army
Come along – x – *Ø/enter the scene/was introduced ++*
(Then feminism **came along** and women all over the world)
2 **Chance** – x – *opportunity*
(You never had the **chance** to see what you really like)
2 Crazy – x – *mentally ill*
(Under special circumstances, like if the mother was **crazy**)
Dark-skinned people - "mørkhudet mennesker" - *minorities*
(**Dark-skinned people** are maybe the most exposed group)
Fancy – x – *extravagant/lavish ++*
(I am truly amazed by all the **fancy** products from different industries)
Hook up to - "hooke opp til" - *connect to*
(You start your working day only by **hooking up to** the company's server)
Get along with– "komme overens med" – *be on good terms with*
(You have to **get along with** your fellow soldiers)
2 **Get away with** – "slippe unna med" – *face no repercussions/Ø*
(Grown people […] expect to **get away with** anything)
Get 5 – x – *go/travel*
(To **get** from one place to another)
Get 6 – x – *to give*
(What **to get** Monica for her birthday)
Get 7 – x – *Ø/achieve*
(You won't **get as far**)
Get 9 – x – *enter*
(Their main concern was to **get to** heaven)
Get back – x – *return*
(When they **got back** they dinner on the table and a clean house greeted them)
Get home – x – *come/arrive at home*
Ex. When he **got** home
3 **Get married** – x – *marriage*
(I have written down arguments for **getting married**)
Get the opportunity – x – *Ø/be able to*
(The more average people want more money to **get the opportunity** to be busier)
Get through – x – *survive*
(In order to **get through** it at all)
3 Get rid of– x – *remove/discard*
(A way of **getting rid of** unwanted elements)
Give [someone] up– x – *give away/have to part with something/one*
(Carrying a child for nine months, and then **give it up**)
Go away – x – *disappear/vanish*
(The weeds will never **go away**)
2 **Go for** – x - *pursue*
(When the people that don't know what job they should **go for**)
Go get [something] – x – *pursue*
(If men want something, they **go get it**)

Hang out – x – *spending time together ++*
(Phoebe and Rachel **hanging out** at their favorite coffee bar)
Has/have got– x – *have*
(It **has got** the features of a drug)
Keep [someone] down – x – *oppress*
(And besides from making people passive, it also **kept them down**)
Keep up– x – *stay at the same level ++*
(Those who can not [afford expensive things], will not be able to **keep up**)
Look – "utseende" – *appearance*
(A pretty **look**)
Make of oneself– x – *become successful ++*
(Few have any idea of what they want to **make of themselves**)
[Have] **off** – x – *Take/have a vacation*
(He hadn't had **a week off** work since he was 20)
Old people – x – *the elderly/senior citizens*
(When we grow old most of us go living in an institution for **old people**)
Snatch – x – *steal/robbery/shoplift*
(Perhaps a successful **snatch** triggered a more serious crime)
Something on [television] – "noe på tv" – *watch television/Ø*
(Neighbors and friends often came to watch when there was **something on**)
2 **Stuck** [somewhere] – x – Ø
(You are **stuck** here because you need a job)
Stupid – "dum" **-** Ø
(We are made passive and **stupid**)
The whole world – x – Ø
(Know what happens in every soap opera in **the whole world**)
In other **terms** – x – in other *words*
(In other **terms**, she fought for women's rights)
2 **Totally**– "helt" – *completely/ vastly different*
(**Totally** different method)
Turn out – x – end/occur/Ø
(Things did not **turn out** exactly as planned)
Up to – x – responsibility
(It is all **up to** the government)
2 **[A] waste [of] time** – x – *Ø/not profitable/worthwhile*
(There is not point in dreaming. It's **a waste of time**)
Whatsoever – x – Ø
(Often nothing is done to prepare the felons rehabilitation **what so ever**)

Connotational
Severely – x – *greatly* (Severely can only work in negative contexts)
(The situation has improved **severely** in recent years)

Synonym/Syntactic
5 tokens, 5 types L1 influence: 0 tokens
Go to great **achievements** – x – go to great *lengths*

(You don't have to go to great **achievements** to break away from it)
Advantage – x – get the advantage/in [his/her/their]advantage ++
(Man has altered many things in order to achieve **advantages**)
Practically good – x – *well, practically speaking/practically Ø*
(If they can do the job **practically good**)
The statement **says** – x – statement Ø
Empty-minded – x – *mindlessly*
(Or more likely surf **empty-minded** on the internet)

Synonym/Invalid
1 token, 1 type L1 influence: 0 tokens
Inacceptance – x – *intolerance*
(When it comes to **inacceptance**)

ASSOCIATIONS
21 tokens, 19 types L1 influence: 1 token

Association/Semantic
9 tokens, 9 types L1 influence: 0 tokens
Bomb – x – *bombard*
(Companies **bombed** them with tempting possebilities)
Change – x – *chance*
(They are not given the **change** to prove the crime was …)
Deliverance – x - *delivery*
(Humans do the **deliverance**, but machines do all the preparations)
Descending – x – *condescending*
(Now, people think these **descending** thoughts [racism] to be dead and gone)
(Censorship in the classic sense is almost non-**exciting**)
Imported – x – *important*
(The most **imported** thing is how)
Physics – x – *physical*
(When it comes to **physics** and/or on the psychic level)
Preacher – x – *priest*
Preserve – x – *protect*
(Now, almost every country has a kind of constitution that **preserves** the inhabitants)
Principal – x – *principle*
(The **principal** of punishing criminals is not outdated)
Transfer – x - *transportation*

Associational/Collocational
2 tokens, 2 types L1 influence: 1 token
Boothes – x – *boxes*
(It puts people into **boothes** which does not excist in real life)
Take an assumption – "å an**ta**" – *make* an assumption

(I may have **taken** wrong assumptions)

Association/Syntactical
8 tokens, 6 types L1 influence: 0 tokens

Dependent – x – *depending/are dependent*
(Because different firms **dependent** on the type of …)
Dismiss – x – *be dismissed*
(When they **dismiss**, they receive a large amount)
Excess – x – *excessive*
(Greed and **excess** consumption)
3 A **high** degree – x – a *higher* degree
(Why should the only ones with high degrees get the job?)
In **compare** to – x – in *comparison* to
Situation comedy – x – *situational* comedy

Association/Invalid
2 tokens, 2 types L1 influence: 0 tokens
Consum – x – *consumption*
(They don't have an enormous **consum**)
Predictment – x – *prediction*
(The **predictment** of marrying is that you want to be together forever)

CORES AND LEXICAL TEDDY BEARS
110 tokens, 16 types L1 influence: 50 tokens

Core/Semantic
1 token, 1 type L1 influence: 0 tokens
Through with – x – *done with/have completed/finished*
(When you are **through with** your service)

Core/Stylistic
109 tokens, 15 types L1 influence: 50 tokens

21 **A lot [of]** – "mye" (not stylistically inappropriate) – *quite a bit/much/several* **t**
(Do it with **a lot of** care/ **A lot of** us today)
8 **Big** – "stor" (not stylistically inappropriate) – *large* **C/t**
(**Big** changes/trend)
5 **Boring** – "kjedelig" (not stylistically inappropriate) – *dull/mundane/meaningless* **t**
(The world would be a **boring** place to live)
14 **Deal with** – x – *cope with/Ø* **t**
(It is difficult to **deal with** situations like death alone)
16 **Get 1** – x – *become/be/Ø* **C/t**

(**Get** better)
22 **Get 2** – "få" – *receive/Ø/find/obtain* **C/t**
(Where they **get** mental help/All they could **get**)
6 **Get 3** – "få" – *have* **C/t**
(Ones who **got** their homes turned upside-down/Money is all he has **got**)
4 **Good** – god (not stylistically inappropriate) – wonderful/positive/ ++ **C**
(How **good** it was)
Great – x – *wonderful/excellent* ++ **C**
(They had a **great** time)
5 **Huge** – stor (not stylistically inappropriate) – *large/vast/overwhelming* **t**
(The difference is **huge)**
Make – lage/(not stylistically inappropriate) – *create* **C**
(You need an idea before you can make something)
Nice – x – *wonderful/lovely/kind* ++ **C**
(In spite of **nice** words)
11 **Really** – virkelig – *quite, truly* **C/t**
(I'm **really** glad)
Right – x – *correct* **C**
(This statement is **right**)
2 **Sad** – "trist" (not stylistically inappropriate) – *tragic* **C**
(I found it surprising and incredibly **sad**)
12 **Thing** – x – Ø **C/t**
(The **thing** about/Most important **thing**)
7 **Very** – "veldig" (not stylistically inappropriate) – Ø **C/t**
(**Very** interesting)

APPENDIX 5

Explanation of numberings for *get*

GET 1 – x – become/be/Ø
Ex. **Get** thinner/ my vision **got** blurry/The word **got** smaller

GET 2 – få – receive/Ø/find/obtain
Ex. Where they **get** mental help/All they could **get**

GET 3 – få – have (DELEXICAL VERB!)
Ex. You **got** cancer

GET 4 – få – develop
Ex. There are many people who **get** bad nerves

GET 5 – x – go/travel
Ex. To **get** from one place to another

GET 6 – å gi – to give
Ex. What **to get** Monica for her birthday

GET 7 – å komme – Ø/achieve
Ex. You won't **get as far**

GET 8 – få (not stylistically inappropriate) – are permitted/allowed
Ex. We **get to** learn

GET 9 – x – enter/arrive
Ex. Their main concern was to **get to** heaven