Abstract

While a great deal of research has shown that native speakers are sensitive to the argument structure of verbs in their language, minimal research has been dedicated to sensitivity to argument structure in second language learners. The thesis aims to investigate L2 sensitivity to argument structure in two proficiency groups (upper secondary school and university). Behavioral data was collected through a task where participants responded to incomplete sentences. This method assesses participants' cloze probability on items following selected verbs. Two approaches to sensitivity were explored. The results displayed a general tendency of cloze probability decreasing in line with proficiency level. The cloze probability of the verbs imply that the upper secondary group are less sensitive to the constraint of the selectional requirements than the university group. While both proficiency groups displayed a sensitivity to argument structure in that they met the selectional requirements of the verbs, the upper secondary group showed more signs of lexical transfer from L1 representations to L2. The university group appear more sensitive to the frequency of typical NP-fillers than the upper secondary group. The cloze values for some of the verbs which appear open in selectional requirements, could suggest that the native speakers have formed some kind of association between the specific verb and the most typical NP.

Acknowledgments

While this process has been challenging, it has at the same time been exciting to see that the things I have learned about language acquisition throughout my studies can in fact be seen in my own work, as indications of evidence in my own results. Learning about something through reading is not nearly as rewarding as learning about something through conducting your own experiments, and finding your own results and indications of how things may work.

This thesis would not have been possible without the help and support of my supervisor, Mila Dimitrova Vulchanova, who has guided me through the process, and co-supervisor, Giosuè Baggio, for helping me with my material. I would also like to thank Sindre and Tonje for helping me conduct my experiments, Rositsa for teaching me how to use the corpus, and all the people who took the time to participate in the study. Thank you to Helena, who offered to read and give feedback on the paper.

I would also like to thank Tonje and Lene, for being the best companions these last two years. Sunniva, Ingvild, my friends in “Lesesal D113”, and everyone else that have been part of this master thesis bubble. Even in the most stressful and frustrating days, our conversations, rants and discussions have always made things better. I think that the time spent with you, either at our desks or at “Masterverksted”, has been essential for the writing process and for my own mental health. Thank you, Turi Marte, for creating the master’s workshop, and for arranging seminars about formatting, writer’s block, EndNote, master of disaster, etc. You created a lifeline for us, and I am very glad you convinced me to join in.

Further, I would like to thank the girls in NTNUI Volleyball D1, for distracting me almost every single day throughout the year, and forcing me out of the master’s thesis bubble. Thank you to my roommate, Torild, for being the ultimate supporter on the home front. You have been the Sam to my Frodo. Finally, I would like to thank my family for always believing in me, and encouraging me to continue. Especially, I would like to thank my sister, Line, for being an inspiration and living proof that even though the process of writing a master thesis is hard, both physically and mentally, one does get through it, and life goes on.

Mona Langeng Hammerås

Table of contents

List of tables

1 Introduction

The structure of a sentence relies on the verb and the argument structure of the verb. As Pinker (2013) argues, since the argument structure of the verb is such an important factor in explaining how a language functions, understanding the acquisition of argument structure becomes an equally important part in explaining how language is acquired. One can argue that the same approach applies to understanding the problems of second language acquisition as well. Several studies have shown that native speakers are sensitive to argument structure, in the form that information inherent to the verb can be used to evaluate upcoming input. Kamide, Altmann, and Haywood (2003, p. 133) suggest that there is "an incremental processor that is able to draw on different sources of information (some non-linguistic) at the earliest possible opportunity to establish the fullest possible interpretation of the input at each moment in time". Is this also the case for second language learners? Can they use this information in the same way, and is it the same information as for a native speaker?

The field of psycholinguistics seeks "to uncover the mental representations and processes through which people produce and understand language" (Garrod, 2006, p. 251). The research done in this thesis is relevant as it sheds light on something that there is little research on, at least in the case of second language acquisition and proficiency, and sensitivity to argument structure in processing. We know that native speakers develop a sensitivity to argument structure, but there is little research on how this is manifested in the second language learner, what it is like compared to a native speaker's sensitivity, and if there is a link between level of proficiency and sensitivity. This thesis can help provide a better understanding for the learning process of children and young adults, when acquiring a second language. Therefore, the thesis at hand looks at how sensitivity to argument structure develops in Norwegian second language learners of English, and how the sensitivity of an L2 learner differs from that of a native speaker. By using incomplete sentences, the study aims to provide an understanding of this by looking at different age groups with different proficiency levels, namely one group from the first year of upper secondary school, and one group from university level (at least one year of introductory classes to English). The two L2 groups are also compared to a group of native speakers.

The results were expected to display a difference between the two proficiency groups regarding sensitivity to argument structure. It was expected that the university group would reflect more or less the same level of constraint on verb argument structure as the native speaker group, and that the upper secondary group would show less sensitivity to argument structure and also show signs of transfer from the L1. These results could be linked to degree of exposure to the language, proficiency levels, and working memory capacity. With the aim if the thesis and these expectations in mind, the thesis will explore and assess two different approaches to sensitivity to argument structure, and what is activated upon recognition of the verb. One type of approach is theoretical, where the idea is that the selectional requirements of a verb are lexically encoded in the entry of verbs, and that this information is activated upon recognition of the verb and used to process and predict upcoming input in sentence processing (Koenig, Mauner, & Bienvenue, 2003). This type of approach relies on verb meaning in the form of abstract representations which specify the number and type of participants that the situation denoted by the verb includes. The other type of approach to argument structure is rooted in computational approaches to language, assuming that arguments are represented at a more concrete level, whereby native speakers over time form associations between a specific verb and the possible fillers for the NP slot, and would typically activate NP-fillers which are most frequently used (e.g., similar to collocations).

The thesis was inspired by the work done by Johnsen (2016) and Reine (2016) in two separate theses, which aimed to find out whether second language users are sensitive to argument structure. The theses were a collaborative project, but focused on separate groups with different proficiency levels – one upper secondary group and one university group. An analysis of both groups was not possible due to the size of the projects. The size of this thesis, however, allowed a study which could look at both groups. While these two theses used eye tracking (from the Visual World Paradigm) as a method, the current thesis has chosen a method which focuses on collecting behavioral data from performance on various tasks, tapping both language and background variables. The main language measure assesses participants' cloze probability on items following selected verbs. This is collected in a task where participants respond to incomplete sentences. Importantly, this method also allows for a qualitative approach to the material, looking at factors such as constraints on the verb arguments, frequency, and tendencies of transfer from L1 verb representations to L2.

2 Theoretical background and previous research

The following chapter will present the theoretical background for the thesis, where the first section will be looking at the lexical semantics of verbs, as entries stored in our mental lexicon, and provide a basis for the concept of sensitivity to argument structure. The following section will focus on what it means to be sensitive to argument structure, and how it is acquired in first and second language users, looking at concepts such as statistical learning and transfer, which are relevant for understanding development of sensitivity to argument structure in second language users.

2.1 The lexical semantics of the verb – what determines the type of argument the verb should take?

First, the field of lexical semantics aims to study word meaning, including the morphemes that might make up a word, and multi-word units, such as idioms (Saeed, 2009). This thesis is concerned with the lexical semantics of verbs, and the information that is encoded in their lexical entries, not considering each morpheme that might build up the verbs that are to be considered.

Our knowledge about verbs, or lexical knowledge about words in general, is stored as entries in our mental lexicon in long term memory (Jackendoff, 2002). Jackendoff (2002) argues that both units smaller and bigger than grammatical words can be stored in this lexicon, meaning that stems and affixes are encoded, along with idioms, such as *kick the bucket.*

The thesis at hand is concerned with the lexical entries of verbs, and the information that is stored in them, or in the words of Koenig et al. (2003), what is *lexically encoded* in the entries of verbs. Linguists do not agree as to exactly what information about the verb is actually encoded in the lexical entries. Levin (1993) presents an argument by Bloomfield (1933), saying that only that which is necessary is lexically encoded, and since syntactic behavior is determined by the meaning of the verb, syntactic properties of the verb is not lexically encoded. The most accepted view is, however, that both syntactic (subcategorization frames) and semantic information (selectional restrictions) is encoded in lexical entries (Friederici & Frisch, 2000). Koenig et al. (2003) are more specifically interested in what participant information is encoded in lexical entries, and in their paper, they present a hypothesis for criteria that determine whether participant information is lexically encoded in the entries of verbs or not. They propose two criteria that together comprise the Lexical Encoding Hypothesis. The first is the Semantic Obligatory Criterion (SOC). This states that "if *r* is an argument participant role of predicate *P,* then any situation that *P* felicitously describes includes the referent of the filler of *r*" (Koenig et al., 2003, p. 72). The next criterion regards specificity. The Semantic Specificity Criterion (SSC) states that "if *r* is an argument participant role of predicate *P* denoted by verb *V,* then *r* is specific to *V* and a restricted class of verbs/events" (Koenig et al., 2003, p. 73). Thus, the Lexical Encoding Hypothesis (LEH) claims that "a participant role is a (semantic) argument of a verb if and only if it satisfies both the SOC and SSC, that is, if its presence is required of all situations described by that verb *and* if it is required of the denotation of only a restricted class of verbs" (Koenig et al., 2003, p. 75). Koenig et al. (2003) hence make a distinction between that which is encoded and that which is not, stating that participant roles which are encoded in the entries are *arguments*, and those that are not are *adjuncts.* They argue that the distinction between an argument and an adjunct is important both for linguistic theorizing and for research on human sentence processing, and that "that the human sentence processing mechanism gives precedence to arguments in building the representation of a sentence" (Koenig et al., 2003, p. 69). The following section will take a closer look at arguments, as well as thematic roles (participant roles) and selectional requirements set by the verb.

The verb determines the structure or form of a sentence because it sets selectional requirements to what constituents it must have or cannot have in its presence (Haegeman & Guéron, 1999). Hence, these selectional requirements of the verb determine what will follow the verb, which is a central factor for this thesis and the method that is used. Depending on the meaning of the verb, or the situation/event or state that it denotes, the sentence the verb heads will require certain elements or participants to be involved. This demand for certain participants to be involved is called the argument structure of the verb (Haegeman & Guéron, 1999).

Participant roles, or thematic roles, are a widely discussed topic amongst linguists in the literature, and therefore there are several views on what thematic roles should be used and what they should be called, and also what problems arise with them. Saeed (2009) and Haegeman (1994) both provide an overview of the theta roles that are most typical in the literature. Some of these are AGENT, PATIENT, THEME, EXPERIENCER, LOCATION and STIMULUS. These theta roles will be described in the following, as done by Haegeman (1994) and Saeed (2009). The AGENT is the participant who initiates the action that the predicate expresses. The PATIENT is the theta role given to the participant which is affected by the action in some way. The THEME is assigned to the participant which undergoes some sort of motion initiated by the action expressed by the verb. The two theta roles PATIENT and THEME can be difficult to distinguish, and is sometimes combined into one THEME role (Haegeman, 1994; Haegeman & Guéron, 1999). The participant which is aware of, but not in control of the action or state expressed by the predicate, or experiences a psychological state, is assigned the EXPERIENCER role. The LOCATION is the place where the event or state denoted by the verb takes place. The final role, STIMULUS, is assigned to the entity which causes an effect of some sort in the experiencer (Haegeman, 1994; Saeed, 2009).

The idea of theta roles can be problematic, a matter which is evident in the literature (as stated above). The list of roles proposed by different authors is long because it is difficult to define roles that are specific enough, but again, not too specific. Theta roles that are very general would mean great variation within the role type; theta roles that are very specific or too narrow would mean that the usefulness of the theta roles would be reduced (Saeed, 2009). Dowty (1991) aims to define theta roles through entailments of the predicate. He proposes different properties of two *prototypes* – the Agent Proto-Role and the Patient Proto-Role, where the properties would allow variation within each prototype (Dowty, 1991; Saeed, 2009). For the purpose of this thesis, it is more productive to use the more delimited theta roles that have been described above. Using these theta-roles will allow a narrower characterization of the verb's argument structure/selectional requirements.

As already stated, verbs have certain requirements when it comes to their arguments and thematic roles. It is assumed that these requirements are part of the lexical information stored in the lexical entries of verbs (Koenig et al., 2003). The *thematic role grid,* or *theta grid,* can illustrate this by showing how many arguments the verb or predicate must take, and what thematic roles may be assigned to the arguments by the verb (Haegeman, 1994; Saeed, 2009). Take the verb *put* as an example (Saeed, 2009, p. 160):

(1) **put** V: <AGENT, THEME, LOCATION>
(2) Mona put the box in the cupboard
(3) The box was put in the cupboard by Mona

By looking at this theta grid, one knows that the verb *put* is a ditransitive verb, taking three arguments that may select the theta roles AGENT, THEME and LOCATION. In sentence (2), *Mona* fulfils the role of AGENT, *the box* the role of THEME, and *in the cupboard* the role of patient. This sentence fulfils the selectional requirements of the verb, as specified by the theta grid. While the theta grid specifies what roles are required by the verb, the event or state that the verb denotes determines how these roles can be realized as arguments. This can be illustrated through alternations.

As discussed by Levin (1993), speakers can make judgements on what alternations the verb allows. They know that even though two verbs are closely related, they may not allow the same alternations. Speakers know that the verbs *spray* and *load* may participate in the locative alternation, and that the two verbs *fill* and *pour*, which are similar, do not allow this alternation. Levin (1993, p. 2) illustrates this in her examples:

(4) Spray
  a. Sharon sprayed water on the plants.
  b. Sharon sprayed the plants with water.

(5) Load
  a. The farmer loaded apples into the cart.
  b. The farmer loaded the cart with apples.

(6) Fill
  a. *Gina filled lemonade into the pitcher.
  b. Gina filled the pitcher with lemonade.

(7) Pour
  a. Carla poured lemonade into the pitcher.
  b. *Carla poured the pitcher with lemonade.

While *spray* and *load* can take a locative alternation, the verbs *fill* and *pour* cannot, even though they are closely related. Levin (1993) argues that it is essentially the meaning of the verb which allows the speaker to make judgements about its behavior, in terms of argument structure and syntactic realization. Verbs that are similar in meaning may be categorized into different classes of verbs which have similar behaviors, according to the type of event or state they denote (Aitchison, 2003; Levin, 1993). This implies that alternations are sensitive to different components of verb meaning. An alternation needs certain meaning components to be part of the meaning of the verb in order to allow the verb to take part in this alternation (Levin, 1993). This causes a web of connections between verbs and the alternations they can take.

2.2 Acquiring a sensitivity to argument structure in L1 and L2

2.2.1 Sensitivity in the L1

This thesis is concerned with sensitivity to argument structure in human sentence processing. As pointed out by Kamide et al. (2003), the human sentence processer works its way through a sentence incrementally. Word by word, it uses information activated upon recognition of a word to interpret this information and predict the upcoming input. The idea of an incremental sentence processor has been supported through several studies, including a study by Altmann and Kamide (1999). Their experiments used participants' eye movements to target objects as a measure, and found results indicating that information activated when the verb is encountered, can be used to constrain the possible upcoming grammatical object. Whilst hearing sentences such as *"the boy will move the cake"* or *"the boy will eat the cake"*, they were shown pictures of a boy, and a cake, along with distractors. Both experiments in the study showed that eye movements to the target object, which in this case was the cake, started after the onset of the word *cake* in the *move* condition, and before the onset in the *eat* condition. The more constraining the selectional requirements of the verb were, the higher would the probability be of looking at the target object before its onset, i.e. before hearing the noun that followed the verb (Altmann & Kamide, 1999). In a study using ERPs (event-related brain potentials), Friederici and Frisch (2000) investigated the use of both verb-specific and argument-specific information in sentence processing. In the case of verb-specific information, they found that when a verb is presented early in a sentence, the upcoming argument is checked against the semantic and syntactic requirements set by the verb. The same was found when argument-specific information was presented early in the sentence; the upcoming verb was checked against the argument-specific information already presented (Friederici & Frisch, 2000). Hence, when a lexical item is recognized, the entry is activated and semantic and syntactic information can be used to predict upcoming input. An interpretation and prediction of upcoming input is based on constraints from semantic, syntactic and real world information (Altmann & Kamide, 1999; Boland, 2005; Kamide et al., 2003; Meints, Plunkett, & Harris, 2008).

With studies like this in mind, one can say that native speakers are sensitive to the argument structure of the verb. This sensitivity means that the verbs selectional requirements, or argument structure, is lexically encoded in the entries of verbs, and when the verb is recognized in sentence processing, the entry is activated and the selectional requirements are used to predict the upcoming input, or the post-verbal argument. This activation is the key for the method of this thesis, as one can assume that the activation of lexically encoded information upon recognition of the verb increases the possibility of it being used to complete the sentence (Koenig et al., 2003).

One important aspect of understanding human sentence processing lies in how we access the mental lexicon. Research has shown that when we process acoustic input, we activate all lexical entries that are compatible with the input in our search for the right item to recognize (Altmann, 1998). Altmann (1998) points out that an item can be recognized before the offset of the word, and one of the factors which determines whether a word is recognized faster is the frequency of the word in the language. This frequency effect has been found in both visual and auditory word recognition (Forster and Chambers, 1973, and Marslen-Wilson, 1987, in Altmann, 1998). This aspect of sentence processing is essential for this thesis in that it may imply a frequency effect in responding to the incomplete sentences. This means that when responding to the incomplete sentences, the recognition of the verb will activate an item which is most typical to serve as the argument of the verb, based on the word's frequency of occurrence in the language. That is, its frequency of occurrence as the NP for the argument slot of the verb.

Now that it has been established that native speakers are sensitive to argument structure, it is essential to look at how children acquire their knowledge of verbs, as it may help provide a better understanding of how second language learners acquire sensitivity. Meints et al. (2008) investigated verb comprehension in young children from 15 months to 3 years old, and asked the question of whether young children link thematic roles to verbs the same way as adults, or if they will allow any participant to be assigned to the given role. They used patient typicality to investigate "whether verb selectional restrictions are applied with our without consideration of real-world knowledge context" (Meints et al., 2008, p. 440). They found that at 18 months, children would allow both typical and atypical verb-patient mappings (eating apples vs. eating houseplants), at 24 months, children's mapping become more narrow, allowing only typical patients, and at age 3, children were able to accept atypical verb-patient mappings, despite their world knowledge implying that it is an atypical pairing (Meints et al., 2008). Meints et al. (2008) hence argue that at age 3, children have developed an adult like mapping, where they are able to integrate their knowledge of verbs and their arguments with their knowledge of thematic roles and the real world.

2.2.2 Sensitivity in the L2

As mentioned in the introduction, this thesis was inspired by the combined work of Johnsen (2016) and Reine (2016), who in their master's theses both investigated sensitivity to argument structure in Norwegian speakers of English. The study used eye-tracking to measure gaze proportions towards a target object, where "the aim was to see whether the level of constraint would affect the gaze behavior by way of increased looks towards the appropriate object for verbs that were more constrained in their argument structure" (Johnsen, 2016, pp. 1-2). Johnsen (2016) looked at a group of participants from upper secondary school, and found that for the most constrained verb category, there was a higher rate of gaze proportions than for the moderate and least constrained categories, suggesting a certain sensitivity to argument structure. For the university group, Reine (2016) argues for a strong association between gaze proportions and the most constrained verb category, and suggests that there is a clear sensitivity to arguments structure. A combined statistical analysis of the data collected by Johnsen and Reine has not been conducted. However, looking at their mean values for gaze proportions towards the target item suggest a higher sensitivity to argument structure in the more proficient group, i.e. the university group. The upper secondary group had a mean value of 0,49112 for proportion of looks in the most constrained verb category (Johnsen, 2016, p. 25). In comparison, the university group had a mean value of 0,60423 in the same category (Reine, 2016, p. 28).

The challenges an L2 learner faces in developing sensitivity to argument structure can in one way be understood through the notion of lexical relativity. With the lexical relativity hypothesis, Stringer (2010) refers to the idea that the meaning of a lexical item is relative to other lexical items in its ambient lexicon, which again means that different languages will lexicalize concepts in different ways. At a lexical level, conceptualization of states and events is to a certain degree specific to each individual language (Stringer, 2010). This can cause difficulties for a learner of a second language.

Stringer (2010) argues that there is lexical transfer from the L1 in the process of second language acquisition. He argues that instead of there being a transfer of parameter settings (e..g, syntax) from the L1 to the L2, L2 errors can be explained by lexical transfer, from the L1 lexical item to its equivalent L2 item. Thus, the L1 lexicon serves as the initial state for the L2 lexical acquisition. The syntax and semantics of a verb will remain unchanged until the speaker has been exposed to an appropriate and sufficient amount of input, which will allow the speaker to reanalyze the item (Stringer, 2010). Stringer (2010) uses the Korean translation of the English verb *put* as an example. The verb *put* can be translated in multiple ways in Korean. Thus, a Korean learner will initially associate the L2 label *put* with one or more of the analogous L1 verbs, making it an interlanguage interpretation. With enough input, this may be altered into the L2 target. Thus, one cannot put an equal sign between items of the L1 and items of the L2. In the beginning of the acquisition process, however, that is in some ways what the L2 learner does, which causes difficulties and mismatches. Comparing the responses of the second language participants to the native speaker responses, and the university group to the upper secondary group, may show cases of this type of transfer from L1 to the L2 in this specific group. Lexical relativity can help provide an understanding of the challenges that learners face in acquiring the argument structure of the L2; what their initial state in learning is like, and how it develops.

All the verbs used in the experiment for this thesis (incomplete sentences) can be transitive, meaning they typically take two arguments in the form of subject and object. They differ, however, in the meaning components that make up the meaning of the verbs, and their selectional requirements. This may in turn differ from the analogous Norwegian verbs, which, as a consequence, may cause difficulties for the less proficient L2 learner. If the L1 lexicon serves as the initial state of the L2 learner, and becomes their initial L2 lexicon, the meaning components for the verbs of the L1 will transfer directly as the meaning components of the L2 verbs, causing the L2 learner to allow the same behavior for these L2 verbs, even though they might only apply to the L1 (Stringer, 2010). As Stringer (2010) points out, the lexicon is relative, and two different languages rarely conceptualize items the same way.

Ellis (2002) argues that, as language users, humans are sensitive to frequencies of words in our language. Through exposure to input, we acquire knowledge about the relative frequencies of different features in our language, for example about the frequency of verbs appearing in transitive or intransitive structures, and of what kinds of subject and object a verb most typically takes (Ellis, 2002). Schmitt and Dunham (1999) found indications suggesting that second language learners who were at a moderately advanced level, had intuitions about the word frequency in their second language. Hence, this knowledge about frequency can also be developed for a second language.

Treffers-Daller and Calude (2015) looked at how frequency of words and patterns in the input can assist L2 learners in adjusting the frequency of which they use these words and patterns in their own L2 output to a level that is similar to that of native speakers. They found evidence showing that L2 learners were to a certain degree able to match native speakers in frequency of use, but that sensitivity to frequency became stronger as proficiency levels became higher. This type of statistical learning is an important tool for L2 learners to overcome generalization and transfer from L1 to L2 (Treffers-Daller & Calude, 2015).

2.3 Working memory and language competence

Lexical information about verbs is stored in long term memory. However, an important aspect of word learning is the process of consolidation and integration of lexical information which is aided by the working memory system. Working memory serves an important role in learning, comprehension and reasoning, and can both store and manipulate information from the outside world and from long term memory/mental lexicon (Vulchanova, Foyn, Nilsen, & Sigmundsson, 2014). Working memory aids the child in developing cognitive abilities and acquiring new skills (Pickering & Gathercole, 2001). Thus, by assessing a child's working memory, one can obtain a better understanding of how well these developments have taken place, and also of the difficulties a child faces in learning and development. The subtests that have been chosen for this thesis are designed to test two of the components of working memory, namely the central executive and the phonological loop (Pickering & Gathercole, 2001). The phonological loop stores verbal information (both spoken and written) temporarily, and is accessed through short term memory (STM) (Vulchanova et al., 2014). It is argued that this component of working memory supports both first and second language acquisition (Pickering & Gathercole, 2001; Vulchanova et al., 2014). The central executive is responsible for controlling the information flow within working memory, and between working memory and long term memory. Forward and backward digit recall, the two subtests chosen for this thesis, tap these two components of working memory. While the forward digit recall task involves significant storage and minimal processing, the backward digit recall task involves significant demands on both parts. Therefore, forward digit recall can be used to measure the phonological loop and short term memory capacity, and backward digit recall can be used to tap the central executive component and measure working memory capacity (Vulchanova et al., 2014). These two subtests will be explained in the description of the methodological design. The results of the working memory tests may provide insight into the differences between the groups, and their challenges in developing argument structure sensitivity.

3 Material and method

This project has had both a quantitative and qualitative approach. The entire project consisted of several tests – one main survey where the participants were to complete incomplete sentences, and four additional tests for additional information on proficiency levels and working memory. These were two online tests (grammar and vocabulary), and two working memory tests (forward digit recall and backward digit recall) from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001). Both the responses to the incomplete sentences and the test on proficiency and working memory have been statistically analyzed independently. However, the material was of such a nature that an analysis on correlations between incomplete sentence responses and additional tests was not possible. Therefore, the discussion will have a qualitative approach to the data. The following sections of the chapter will provide a closer description of the participant groups, the procedure, and the analysis that was conducted on the material.

3.1 Participants

The study consisted of three groups of participants, one group of native speakers of English, and two groups who are native speakers of Norwegian, and have English as their second language (L2). For the two L2 groups, the aim was to have 25 to 30 participants in each group. Due to some participants deciding to pull out of the study at the last minute, and some not being available for the last round, it was difficult to obtain that number. The purpose of having two groups of participants was to gather material from groups with different levels of proficiency, making it possible to see if there was a correlation between an increasing proficiency level and how sensitivity to argument structure develops.

The first group was a class from the first year of upper secondary school (10 years of education), with the age of 15-16 years. The final number of participant for this group was 24 (14 were male and 10 were female). The participants were recruited through contact with an upper secondary school in the area. The second group was comprised of university students who had studied English for at least one year (at the least finished the introductory classes for English). How many years they had studied English varied within the group. The final number of participants for this group was 26 (7 were male and 19 were female). Mean age was 23 years old (ranging from 20 to 31). The university group had an average of 16,4 years of school/education (ranging from 14 to 19,5). All participant for both groups had Norwegian as their native language.

The study was approved by NSD (Personvernombudet), and in line with the rules stated by the NSD, both groups signed a consent form which stated what the study involved, and their rights to withdraw from participating at any time during the process (Appendix F). Due to the age of the upper secondary group, the consent forms had to be signed by the participants' parents. Since the participants were to be anonymous, each participant was assigned a participant number, to ensure that the information collected through the different tests was connected to the same participant.

As mentioned, a native speaker group was also part of the study. They were, however, not participants of this thesis, but participants in the preparatory stage for the work done by Johnsen (2016) and Reine (2016). The participants were native speakers of English from a British university, and made up a group of 105 participants. Mean age was 19,75 (ranging from 18 to 45; 86% female and 14% male) (Johnsen, 2016; Reine, 2016). The material shows only 101 responses for each verb, implying that some of the responses were excluded from the material. Since the native speakers were not participants of this thesis, there is no additional material from online tests (vocabulary and grammar) and working memory for this group.

3.2 Procedure

3.2.1 Online-tests

All participants for both L2 groups had to complete two online tests on their personal computers, one for vocabulary and one for grammar. The vocabulary test[1] was developed at the Center for Reading Research at Ghent University in Belgium. The participants were presented with one word on the screen at a time (100 words in total), where the word could be either an existing English word, or a non-sense word/non-existing English word. They were to click yes if they knew and understood the word, and no if they did not. The score was given in percentage, and indicated how many existing English words they knew, how many non-existing words they clicked yes to, and the total percentage of words that they knew. The results from the test were checked and written down by the experimenter.

[1] Site for vocabulary test http://vocabulary.ugent.be/

In the grammar test[2], the participants were presented with a total of 15 sentences that were incomplete. For each incomplete sentence, they were given four alternatives to choose from to complete the sentence. They were to choose the alternative they felt was best to complete the sentence. They were given a score based on a scoring system called CEF, which ranges from A1 to C2, A1 being the lowest and C2 being the highest (A1, A2, B1, B2, C1, C2). In order to put these results into statistics, A1-C2 were later replaced with numbers from 1 to 6, where A1 was 1 and C2 was 6.

3.2.2 Working Memory

In addition to the online tests, two subtests from the Working Memory Test Battery for Children, as described by Pickering and Gathercole (2001), were conducted. The WMTB-C makes it possible to do a broad assessment of children's working memory capacities through a number of tests, or subtests. These subtests are constructed to obtain information about the three main components of working memory – the central executive, the phonological loop and the visuo-spatial sketchpad. This test battery has been designed in a way that makes it appropriate to use with children who are aged between 5 and 15, but the different subtests have been used in both children and adults, since the principle is the same (Pickering & Gathercole, 2001).

For this study, only two of the subtests were used, namely forward digit recall and backward digit recall. In forward digit recall, tapping the phonological loop and short term memory (STM), the researcher calls out a list of numbers (from 0 to 9) from the WMTB-sheet (Appendix E), and the participant is asked to listen, recall and name the numbers in the same order back to the researcher. Each block has six sequences of numbers, and for every block the participant is able to complete, the level of difficulty is increased by adding one more number to each sequence, starting with one number in the first block, up to nine numbers in the last block. When the participant has correctly repeated four number sequences in one block, the researcher moves on to the next block. If the participant makes a mistake in three sequences within one block, the trial is stopped.

In backward digit recall, tapping the central executive component of working memory, the researcher calls out a list of numbers (from 0 to 9) from the WMTB-sheet, and the participant is asked to listen, recall and name the number sequence in the opposite order of what the researcher said. The rule for moving on to the next block, and for stopping the trial is

[2] Site for grammar test http://www.examenglish.com/leveltest/grammar_level_test.htm

the same as for forward digit recall. The increase in level of difficulty is also the same, but the first block has two numbers in one sequence, and the last block has seven numbers in one sequence. Both tests give two scores each, one for number of trials that were repeated correctly (trials correct) and one for how many blocks the participant was able to complete (span).

The purpose of including forward and backward digit recall was to see how the groups differ in working memory capacity, and to see how it might shed light on some of the differences that can be seen in performance on the incomplete sentences.

3.2.3 Incomplete sentences

The study used incomplete sentences which had a subject and a verb, but where the NP-filler for the *patient/theme* (or object) argument slot was missing. The participants were asked to fill in and complete the sentences with the first thing that came to mind. They were also encouraged not to overthink their answers and avoid long phrases. In addition, they were asked to state whether they were sure or unsure that their answer was grammatically correct, or if they did not understand the word or sentence. The incomplete sentences that were used were taken from work done in the preparatory stage for the theses of Johnsen (2016) and Reine (2016). Before completing the sentences, the participants were asked to fill out some basic information about themselves – gender, when they were born (month/year), how many years of school and education they had completed, and whether Norwegian was their native language. The background questions and the full list of incomplete sentences can be found in Appendix D.

In an experiment to find behavioral evidence for one of the criteria (SSC) in the lexical encoding hypothesis, Koenig et al. (2003) "assumed that participant information that is lexically encoded is retrieved upon recognition of a word. Because this information is activated, it is more likely to be used to continue a sentence" (2003, p. 82). This assumption was relevant in this study as well. The responses can be assumed to reflect the lexically encoded information about a verb and its argument structure, and thus, also reflect the speaker's sensitivity to argument structure. The responses from the native speakers were included as material to serve as a third group and a "golden standard", representing the target for the second language groups. An analysis of the native speaker responses was relevant for a better understanding of their sensitivity to argument structure, and to have something to compare the L2 responses to. The responses from the two groups of L2 speakers provided a measure of their sensitivity that could be compared to that of the native speaker group, and compared to each other, hopefully showing a link between level of proficiency and sensitivity. As mentioned in the introduction, this thesis explores two approaches to sensitivity. In the theoretical approach, considering the selectional restrictions of the verbs being lexically encoded in the entries, the performance on the incomplete sentences would reflect sensitivity if the responses were correct according to what argument the verb must take, i.e. if it matches that argument structure which is lexically encoded in the entry of the verb. Looking at the performance on the incomplete sentences with a computational approach, the responses in all three groups could reflect the participants' sensitivity to frequency in the language; meaning the frequency of the NP-fillers and what is most typically selected as the argument of the verb.

Admire was not included in the material that was used in the analysis, even though it was part of the incomplete sentences for the university and upper secondary group. This was because the verb was not part of the list of incomplete sentences for the native speaker group. This was also the case for the verb *itch,* but since the responses from the two L2 groups show some interesting tendencies regarding lexical transfer, this verb was included in the analysis.

3.3 Analysis

3.3.1 Describing the lexical semantics of the verbs

Some verbs are more constrained than others when it comes to their requirements in argument structure. The meaning of a verb puts selectional restrictions on the category of the object slot, and thus requires certain selectional feature on the object. The theta grid requires a certain type of thematic role for its final argument. To see whether participants and the native speakers were sensitive to these constraints, the selectional restrictions of all verbs used in the survey were described and categorized. For instance, the verb *milk* requires a [PATIENT] argument, and will most likely select an NP that is animate, non-human, such as a mammal (for example *cow).* The NP-fillers given by the participants must reflect these restrictions that constrain the argument structure of the verb. We can assume that the use of NP-fillers with the correct selectional features reflects a sensitivity to the lexical semantics of the verb.

3.3.2 Analyzing performance on incomplete sentences

The main data from this study was collected by having participants in two different groups fill in the blanks for incomplete sentences. As mentioned earlier, in addition to these two groups, similar data from a large group of native speakers of English was also included as material in the thesis (Johnsen, 2016; Reine, 2016). The intention with including the native speaker responses was to use it as a “golden standard”, or a control group, that the university and upper secondary group could be compared to. It was therefore important to analyze this material as well, in addition to the responses from the second language users, in order to obtain a better understanding of the native speaker responses and their sensitivity to argument structure. The analysis involved judging and counting unique answers, and categorizing the verbs according to level of constraint.

The first step of the analysis involved judging and counting the number of unique answers for each verb. When judging and counting unique answers amongst the responses given in the incomplete sentences, only heads of noun phrases were counted (in the case of NPs; if a response was a PP or a subordinate clause, the unique answer was based on the whole label). Determiners and number (singular or plural) was ignored, so that *a shirt, the shirt, his/her shirt* were judged as the same response, as well as *her boyfriend’s shirt* and *her black shirt.* In the case of responses like *room, bedroom,* and *living room,* all three responses would be counted as the same unique response – *room.* Some of the NPs could have denoted different senses, but were still counted as one unique answer, because there was no way of knowing which sense of the head noun the participant was referring to. In the incomplete sentence *“Susan typed…”*, some participants responded with the noun phrase *“a/the letter”.* In this case, one cannot know if the sense of the word intended by the participant was a written message, or a written symbol that represents a sound in a language. Because both senses of the word match the selectional restrictions of the verb, there was no reason to distinguish between the responses.

In the second part of the analysis, the verbs used in the incomplete sentences were categorized into three different classes according to level of constraint. Originally, two different approaches were used. One was based on number of unique responses, and was adopted from the preparatory stage in the work done by Johnsen (2016) and Reine (2016). The other approach is called the cloze method, and was used by Block and Baldwin (2010) when they examined context-based word prediction. The first approach, used by Johnsen (2016) and Reine (2016), distributed the verbs into categories of most, moderately and least constrained. The most constrained category included verbs where three or fewer unique answers constituted at least 55% of the participants’ answers. The moderately constrained category included verbs that had four to six unique answers constituting at least 55% of the responses. The final category, least constrained, included verbs where it took seven or more unique answers to make up 55% of the responses. The second approach, the cloze method, allows the researcher to measure the *cloze probability* of an incomplete sentence, meaning "the proportion of participants who give a certain response to finish an incomplete sentence" (Block & Baldwin, 2010, p. 666). The values used by Block and Baldwin (2010) for cloze probability were low cloze, defined as 0%-33%, medium cloze, 34%-66%, and high cloze, 67%-100%. For this study, this meant that if one unique answer constituted 70% of the responses for an incomplete sentence, the incomplete sentence or verb had a high cloze probability. This cloze probability is thought to reflect the constraint of the verbs argument structure and selectional requirements. For both approaches, the percentage of responses was taken from the whole list of responses, including the once that did not fit the argument structure descriptions for each verb. A comparison of the two approaches to see which one is most useful would have been an interesting contribution to the study, and while the original idea was to use both approaches in the analysis and discussion, the approach used by Johnsen (2016) and Reine (2016) was not included due to the limitations of the scope of the thesis.

The responses from the university group and the upper secondary group were analyzed the same way as for the native speaker group, including judging unique answers and distributing the verbs according to level of constraint. The distribution of the verbs was done using the same cloze method as for the native speakers. In addition to counting unique answers, the number of "*unsure"* and *"do not know the word"* responses were counted for each verb, not considering whether they completed the sentence or what that response was. Finally, if a participant did not respond to a sentence, this was reported as *blank.* Number of *blank* responses for each verb was also counted.

Distributing the verbs for each group according to cloze probability (high, medium, low) can illustrate how sensitivity to the constraint of the verb's argument structure develops in line with proficiency. Furthermore, the cloze probabilities of the verbs and the most frequent NP-fillers for each group can give an indication of the L2 groups sensitivity to frequency, compared to the native speaker group.

Comparing the cloze probability of the verbs to its selectional requirements will also be an important part of the discussion. The descriptions of the selectional restrictions of the verbs show that some verbs are more constrained than others when it comes to what they can take as an argument, and select as NP-filler in the object slot. Looking at the results with a computational approach, the levels of constraint and frequency of NP-fillers for each verb might indicate a different constraint on the verb's argument structure.

3.3.3 Analyzing participant information

The results from the university and upper secondary groups, including both online tests (vocabulary and grammar) and the two working memory tests (forward and backward digit recall), were filled into Microsoft Excel and statistically analyzed in the statistical computing program R[3]. Mean value and standard deviation for each test in both groups was calculated, and because of the nature of the data, the non-parametric Wilcoxen test was used to see whether the differences between the groups could be considered real, or if they were by chance. The quantitative results from these tests can be used to show the differences in proficiency levels and working memory capacity between the groups, as well as within the groups. The material from the incomplete sentences (level of constraint, frequency of NP-fillers, total number of unique answers for each verb, etc.) were of such nature that a quantitative analysis of correlations between the participants' responses and their proficiency levels and working memory was impossible. Instead, the results on proficiency levels and working memory can be used as possible explanations for the features or tendencies that can be seen in the performance on the incomplete sentences.

3.3.4 Corpora

A corpus search was conducted, using the NOW Corpus (News on the Web)[4]. The corpus was used when exploring the computational approach, to provide additional information in the analysis, looking at the high cloze probability verbs and their most frequent NP-fillers, to see if the frequency of occurrence in the language can justify the differences in cloze probability and the openness of the selectional restrictions of the verbs. The NOW Corpus reflects current usage and frequency, and was therefore used to corroborate the native speaker data.

The search was set to include 1000 results (#HITS). For each verb that was part of the corpus search, six queries with different settings were carried out to check for the frequency of the verb with the specific NP-filler compared to the frequency of the verb with any noun. The searches were as follows: the occurrence of a verb with any noun (e.g. brushed_v NOUN), a verb with one word and any noun (e.g. brushed_v * NOUN), a verb with two

[3] R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[4]NOW Corpus (News in the Web): http://corpus.byu.edu/now/

words and any noun (e.g. brushed_v * * NOUN), a verb with an NP-filler (e.g. brushed_v hair), a verb with one word and an NP-filler (e.g. brushed_v * hair), and a verb with two words and an NP-filler (e.g. brushed_v * * hair). The percentage was calculated to show how much of the total occurrence (any noun) was made up of the specific NP-filler for the verb. It was also noted whether the NP-filler was listed as the most frequent NP to occur with the verb.

4 Results

4.1 Selectional requirements of the verbs

The following section shows the descriptions of the selectional requirements for each verb. The descriptions provide an abstract measure of what the verbs are most likely to take as their argument. Some of the verbs appear to be more constrained that others in what they can select. Verbs like *milk, brush,* and *extinguish* appear to be very restricted. Verbs like *remove, kick, read, drop,* and *wash* appear to be quite open. The descriptions will be addressed in the discussion, with relation to the type of responses the participants have given and if the responses match these requirements, and to how constrained the verbs appear with regards to the cloze probabilities of the groups.

She annoyed – [EXPERIENCER] argument (Levin, 1993; Saeed, 2009), assumed to select an NP for the argument slot which is of human character. The argument could also be an animal-denoting NP.

The board approved – [STIMULUS/THEME] argument (Levin, 1993), will select something that needs approval, like a project, action, event, or suggestion (typically an abstract entity).

The policeman arrested – [THEME] argument, assumed to select a human entity, individual person or group.

The dog ate – [PATIENT] argument, most likely to select anything that can be characterized as solid food.

Mila boiled – [PATIENT] argument, most likely a physical entity that is edible, either solid (potatoes) or liquid (water). *Boil* could also be used as a causative construction, and may therefore select *kettle* as its argument.

Robert borrowed – [THEME] argument, most likely to select a physical entity in the argument slot. There is little constraint on what the entity can be.

She bought – [THEME] argument, most likely a concrete, non-human entity, anything that can be obtained through money.

He broke – [PATIENT] argument, most likely a physical entity, or part of the human body.

She brushed – [LOCATION/PATIENT] argument, an appropriate body part (human or animal) or surface that needs to be taken care of or groomed (Levin, 1993).

Nancy buckled – [THEME] argument, a concrete entity that can be fastened with a buckling mechanism.

Sarah buried – [THEME] argument, most likely a physical entity, human (*the body,* or *her face (in her hands)* or animal. Could also be something abstract, like secrets or lies.

They carried – [THEME] argument, most likely to select an NP that is concrete, a physical entity, that can be lifted and moved around.

He changed – [THEME] argument, not very restricted as to what it can select as an argument, as long as it is something that can be changed.

He collected – [THEME] argument, can select almost anything as an argument for the argument slot, but most likely something that is typical to collect as a hobby (*stamps*).

Kevin cut – [PATIENT] argument, can select any physical entity that is possible to make incision into or split by cutting.

Harry decorated – [LOCATION/PATIENT] argument, will most likely select an NP for its argument slot that is a place/location (house, room, etc.) or any physical entity that is normal to decorate (*Christmas tree, cake*, etc.).

She delivered – [THEME] argument, can select almost any NP as an argument for its argument slot. Most typically a physical entity, like a package or parcel. Something that can be passed on to someone else.

She dropped – [THEME] argument, can select almost any physical entity for the argument slot, as long as it is the right size and hence can be lifted/carried and then dropped. Can also select an abstract entity, like *drop the bomb, drop a hint.*

He emptied – [PATIENT/LOCATION] argument(Levin, 1993), an entity that is a container of some sort, that can contain something (*the trash, his mind*, etc.).

Lily entertained – [EXPERIENCER] argument, the NP would be something animate, most likely a human.

I examined – [THEME/LOCATION] argument, can select almost any NP to fill the argument slot. Most typically a document or something that needs attention.

Emma extinguished – [PATIENT] argument, most likely something concrete that can be stopped from burning – something that is in flames.

David filled – [PATIENT/LOCATION] argument, any entity that can contain something, container of some sort.

James fired – [PATIENT] argument, any human being that is employed and in the position to be dismissed from their position. Can also choose a [THEME] NP for the argument slot, an entity that can be sent off/discharged (*arrow*) or can discharge another entity at a great speed (*gun, cannon*).

The woman fried – [PATIENT] argument, assumed to select an NP for the argument slot that is solid and edible.
She frightened – [EXPERIENCER] argument, assumed to select something animate, either human or animal.
Linda ground – [PATIENT] argument, the NP in the argument slot will most likely be a substance that can be crushed into fine powder, or something that can be shaped by rubbing it on to a hard surface. Most likely something edible, like a spice, herb, or a seed of some sort.
Olivia guarded – [THEME/LOCATION] argument, not very restricted and can select any physical object or possession, like a house, a human being or area.
Karen hammered – [THEME] argument. This verb will most likely choose an NP for the argument slot that can be knocked on by a hammer, like a nail.
He hunted – [THEME] argument, most likely an NP that is animate, most likely an animal that is normal to hunt, or can be sought (Levin, 1993).
Jennifer impressed – [EXPERIENCER] argument, any NP that is a human and can experience the change of psychological state.
She ironed – [PATIENT] argument, assumed to pick an NP for the argument slot that is a piece of clothing.
Chris judged – [THEME/EXPERIENCER] argument, most likely to select as an argument a NP that is human, or an event of some sort.
She juggled – [THEME] argument, selects an argument that is either a physical entity that can be thrown around in the air, or an NP that represents an abstract concept, for example *three jobs, work and family, etc.;,* things that are time consuming.
She kicked – [THEME/LOCATION] argument (Levin, 1993; Saeed, 2009), can select almost any physical entity, human or non-human, that can be kicked using your leg and foot.
They knitted – [PATIENT] argument, takes an argument for the argument slot that is a piece of clothing or product that is handmade/ can be knitted.
She knocked – [LOCATION] argument (Levin, 1993; Saeed, 2009), most likely to select a PP for the argument slot. Typically an entity or location that can be knocked on to get someone's attention.
Paul landed – [THEME] argument, will most likely select as an argument a type of vehicle that can fly (plane, helicopter), or an abstract concept that is difficult to get, for example a job or a promotion.

He licked – [LOCATION] argument, most likely to select an argument that is edible, i.e. a lollipop or an ice cream, or something that is typical to lick without it being edible (for example stamps or your fingers). Must be a physical entity.

Andrew lit – [PATIENT/LOCATION] argument, (change of state verb), something that can create fire or be in flames, like a candle or a fire in itself.

Mary loaded – [PATIENT/LOCATION] argument, most likely to select as an argument an entity that can contain something and be filled up.

She married – [PATIENT] argument, selects an argument that is a human being.

Mark measured – [THEME/EXPERIENCER/LOCATION] argument, can select as an argument the attribute of an entity (length, weight, etc.), or entities/objects/concepts that have values that can be measured.

He memorized – [THEME] argument, most likely something that is written or organized in a list of some sort or in a specific order, abstract or concrete.

George mended – [PATIENT] argument, most likely to select as an argument any entity or artifact that can be broken and fixed again. More likely to be an artifact than animate.

Laura milked – [PATIENT] argument, most likely an animal (mammal) that produces milk.

Edward missed – [STIMULUS] argument, as a psychological state, can select as its argument an NP that refers to a human, a human relation, activity, or anything that one can miss as a psychological state. Could also be a [THEME] argument for a verb describing a spatial relation (Levin, 1993). It can select as an argument an event/appointment/chance/opportunity, something that has a time limit (*the bus*), or something that can serve as a target, both in the literal sense and as an abstract target, like missing *the point.*

He moved – [THEME] argument, can select almost any physical entity which is capable of changing location, assuming the verb takes the transitive causative alternation. The verb can also be used intransitively, denoting the event of relocating oneself or the subject.

He obeyed – [THEME/STIMULUS] argument, can select as its argument an NP that is human (*his boss)*, a command, or an instruction.

Thomas ordered – [THEME] argument, will most likely select something edible (solid and liquid) as an argument for the argument slot. Could also be something that can be ordered through online shopping.

He owned – [THEME] argument, the verb can take as its argument anything that can be claimed ownership to, for example a house, a piece of land, or a pet.

Bob painted – [PATIENT] argument, can select as an argument something that can be created through painting (a landscape, a portrait), OR [THEME/LOCATION] argument, selecting as an argument something that can be painted on (a wall, a house).

Roger pinched – [LOCATION] argument, most likely to select an argument that is a human or a body-part. Could also be an object that can be stolen [THEME], like a chocolate bar.

Ruth planted – [THEME] argument, most likely to select something that can grow, like a plant, flower, or tree.

She played – [THEME] argument, will most likely select an NP that is a musical instrument of some sort, or a type of game or sport.

She pushed – [THEME] argument, can select almost anything for the argument slot. However, the entity should be capable of moving, as a response to the force administered by the agent.

She read – [THEME] argument, can select an argument that is any form of literature or written text.

The man reduced – [PATIENT] argument, will select as an argument something where the value of an attribute can be made smaller/reduced, like the weight, a price, or the speed of a car.

Charles removed – [THEME] argument, can select almost any entity to be the object in the argument slot. Any entity that can be removed from a location.

He rubbed – [THEME/LOCATION] argument, most likely to select as an argument a part of the body.

Ronald scratched – [LOCATION] argument, can select an animal or human body part as an argument, OR [PATIENT] argument, can select almost anything as an argument (in the sense of scratching something and leaving a mark, like scratching a car).

John served – [THEME] argument, can select as an argument a human being or group of human beings, an organization (providing a service for someone), or something edible (both solid and liquid).

He sipped – [THEME] argument, will select as its argument a drinkable liquid.

He squeezed – [PATIENT/THEME] argument, will select something that contains a substance/liquid that can be squeezed out, like a lemon or a bottle. Can also select as an argument a body part, or an object that can change shape temporarily by being squeezed, like a ball or a sponge.

They stroked – [THEME] argument, most likely to select as an argument an animal/pet, or a human body part.

She threw – [THEME] argument, can select almost any entity/NP as its argument that is capable of moving through the air as a result of being thrown.

Hannah tied – [THEME] argument, will select something that can be fastened together, or made into a knot, for example a shoelace, hair, or a rope.

He tightened – [PATIENT/THEME] argument, selects something that can experience the change of state in becoming more firm or less loose, like a rope, a grip, a belt, etc.

Brenda twisted – [PATIENT] argument, most likely to select an argument that is a body part, more specifically a joint body part, like an ankle, or a knee (this would be in the sense of damaging a part of the body (Levin, 1993). This verb can also select an entity/concept whose shape can be changed (both literally and figuratively) (Levin, 1993), for example *her hair* or *her words*.

Susan typed – [PATIENT] argument, will select anything that is or can be written using a machine (typewriter or PC).

Sandra unfastened – [PATIENT] argument, will most likely select something that has a buckling mechanism or can be loosened, like a seatbelt.

She unlocked – [PATIENT] argument, will select any entity with a locking mechanism.

Lisa visited – [THEME/LOCATION] argument, either a place/location or a human being/group of human beings.

He washed – [PATIENT/THEME] argument, can select almost any entity as its argument, but most typical are items of clothing, dishes or body parts.

The crowd watched – [STIMULUS] argument, can select almost any event or situation as its argument, or anything that can be perceived (Levin, 1993).

He watered – [PATIENT/LOCATION] argument, most likely to select a plant/something that grows, or a place where something grows, like a garden.

He wore – [THEME] argument, any type of piece of clothing or something that can be worn as a piece of clothing.

They worshipped – [THEME/STIMULUS] argument, most likely an argument that is connected to some religious belief, but can also be something (human/concrete/abstract) that can be looked up to.

4.2 Differences between and within the groups

The online tests (vocabulary and grammar) and the working memory tests were conducted to investigate how the groups differed in proficiency and working memory capacity. The results show that there are differences in both proficiency levels and working memory capacity between the groups, and that there is some variation within the groups as well. The mean value, standard deviation, and p-value (Wilcoxen test) for each test can be seen in table 1.

Table 1: Mean value, standard deviation, and p-value for tests on vocabulary, grammar, and working memory.

	Mean value, university	Standard deviation, university	Mean value, upper secondary	Standard deviation, upper secondary	p-value (Wilcoxen)
WMFTC	36.27	5.41	32.13	4.45	W = 180.5, p-value = 0.01072
WMFS	6.00	0.94	5.25	0.74	W = 177, p-value = 0.005075
WMBTC	23.92	6.90	19.29	5.81	W = 187, p-value = 0.01538
WMBS	5.12	1.24	4.21	1.02	W = 179, p-value = 0.007652
Voc. total	57.46	10.70	35.96	13.27	W = 65, p-value = 1.659e-06
Voc. existing	60.38	12.31	45.04	15.35	W = 143, p-value = 0.001056
Voc. non-existing	3.04	5.44	9.08	9.10	W = 445, p-value = 0.00655
Grammar	5.96	0.20	4.92	1.14	W = 137.5, p-value = 2.89e-05

The Wilcoxen test for each of the tests gave a p-value of <0,05, meaning that the probability that the observed differences between the groups arise by chance is small. The results show that the participants in the university group have more developed working memory capacities, and that they are more proficient in vocabulary and grammar. The results for the vocabulary test show that the upper secondary group were more prone to press "yes" for words that were non-existing words in English, indicating that they were more unsure of their vocabulary knowledge. The standard deviation values for the working memory tests imply that there is slightly more variation in scores within the university group than within the upper secondary group. For the grammar and vocabulary tests, it appears to be the other way around – there is more variation in the upper secondary group. Even though the data collected from the incomplete sentences were of such nature that a quantitative analysis of correlations between these results and the performance on incomplete sentences was not possible, the differences between the groups may serve to justify some of the differences that can be found when analyzing the performance on the incomplete sentences.

4.3 Performance on incomplete sentences

Table 2:Mean values for total number of unique answers, "unsure", "do not know the word", and "blank" responses.

	Total u.a., native speakers	Total u.a., university group	Unsure	Do not know	BLANK	Total u.a., upper secondary group	Unsure	Do not know	BLANK
Mean high	12.30	4.86	1.86	0.00	0.00	6.67	4.00	0.00	0.00
Mean medium	20.35	9.59	1.30	0.11	0.07	8.87	2.40	1.53	1.73
Mean low	32.50	15.07	0.76	0.24	0.18	13.42	3.18	2.90	3.11

Table 2 gives an overview of mean values in each constraint category. For the university group and the upper secondary group there is a mean value for total number of unique answers (u.a) per verb, a mean value for how many participants per verb reported that they were unsure if the sentence was grammatically correct, a mean value for how many reported that they did not know or understand the word, and a mean value for how many gave a blank response per verb. Since the material for the native speaker group was not collected in this thesis, only the mean value for total number of unique answers per verb could be included. Values for each verb is displayed in a table in Appendix B.

The values found in table 2 show that, for all three groups, the mean value for total number of unique answers per verb increases as level of constraint decreases. For the university group and upper secondary group, one can see that the upper secondary group have higher mean values for "unsure", "do not know", and BLANK, indicating a lower proficiency level than the university group. These results reflect the differences that can be seen between the two groups in the results from the tests on vocabulary and grammar.

The following will present tables showing the distribution of verbs according to level of constraint for each group in the high and medium cloze classes. The table for the low cloze verbs can be found in Appendix A. The table for high cloze verbs (table 3), and the two tables for medium (table 4) and low cloze, are organized differently, due to the size of the two latter categories. Even though the verbs have been categorized in classes according to high, medium, and low cloze probability, it is important to emphasize that the verbs show a continuum of cloze probability, not clear cut-offs forming distinctive classes. In addition to displaying the cloze values of the verbs, the most frequent NP-filler or unique answer is also listed.

Looking at the most frequent NP-fillers or unique answers for all three groups (see tables 3 and 4 below, and Appendix A), one can see that most of the responses fulfill the selectional requirements set by the verbs. For the native speaker group, all the unique answers fulfill the selectional requirements, as described according to the transitive realizations of the verbs. Looking only at the most frequent unique answers, there are some verbs that appear to have been realized as intransitive verbs. For the university group, these are *move, itch* (can only be intransitive), and *type.* For the upper secondary group, only *examine* and *move* have unique answers that indicate an intransitive treatment of the verb. For both the university group and the upper secondary group, the cloze values for these verbs are low. Even though these verbs have been realized intransitively, it does not indicate insensitivity. This is a behavioral study, and we cannot control how the participants interpret the verbs.

Even though the general tendency for the L2 groups is that their responses match the selectional restrictions set by the verb, a closer look at the responses for some of the verbs indicate that there are tendencies of some participants not fulfilling these requirements. These are *itch, load, serve,* and *examine,* and will be further addressed in the discussion, with regards to transfer and exposure to input. An overview of all the unique answers for these verbs can be found in Appendix C.

Table 3: High cloze probability verbs (100-67%), showing unique answer and cloze probability for each group.

High	Unique answer	Native speaker group	University group	Upper secondary group
Extinguished	fire	0.95	0.85	
Milked	cow	0.93	0.92	0.88
Hammered	nail	0.86	0.81	
Watered	plant	0.80	0.77	
Kicked	ball	0.75		
Unlocked	door	0.75	0.85	
Unfastened	seatbelt	0.74	0.85	
Read	book	0.71	0.77	0.75
Knocked	(on the) door	0.69		
Brushed	hair	0.67	0.65	
Ate	food			0.67

Table 3 illustrates the distribution of high cloze verbs for all three groups. The verbs are arranged after cloze probability value for the native speaker group. The verbs where the groups overlap have the same unique answer. The native speaker group have ten verbs in this category, the university group have eight, and the upper secondary group have three. For the university group, *brushed* was placed in the high cloze category, even though the cloze probability was below 67%. *Brushed* had a cloze value of 63% in the upper secondary group, but since their most frequent NP-filler was *teeth* instead of *hair*, the verb was placed in the medium cloze category. Some of the verbs show a high cloze value, even though they appear open in selectional restrictions. *Kicked* had a cloze probability of 75% for the native speaker group, even though the selectional restrictions of the verb display a very open verb. All the verbs for the university group match the native speaker group. The upper secondary group only have two verbs that match the native speaker group, and one verb *(ate food)* that does not match either of the other groups.

Looking at the high cloze probability verbs for the native speaker group and the university group, one can see that the university group only lack two of the verbs from the native speaker list, namely *kicked* and *knocked.* These are listed in the medium cloze category (table 4 below), but still have a high cloze value – *kicked* has a cloze probability of 54%, and *knocked* has 50%. The upper secondary group also have *kicked* (46%) and *knocked* (38%) in the medium cloze category, but display an even lower cloze value than the university group.

The upper secondary group lack eight of the verbs listed in the high cloze category for the native speakers, including *kicked* and *knocked.* Some of these are listed in the medium cloze category, including *unfastened* with 63%, *brushed* with 63%, *kicked* and *unlocked* with *42%,* and *knocked* with 38%. The remaining verbs that were ranked as high cloze probability verbs in the native speaker group, were listed as low cloze probability verbs in the upper secondary group. These were *watered* with 33%, *extinguished* with 29%, and *hammered* with 13% (see Appendix A for table on low cloze probability verbs). These results indicate a decrease in cloze probability in line with proficiency levels.

While cloze probability is around the same level in all three groups for several verbs, there is clear tendency for the university group and the native speaker group to be around the same level, and for the upper secondary group to show a lower cloze probability compared to the native speakers and the upper secondary group. This is especially the case for verbs that were high cloze for the native speaker group and the university group.

There is more variation in cloze probability across the groups in the medium cloze category, but the general tendency is still the same – the university group are closer and sometimes similar to the native speaker group in cloze probability, and the upper secondary group seem to display cloze values which are lower than both the native speakers and the university students. Some examples are *worshipped,* where the frequency was 61% for the native speakers, 62% for the university group, and 42% for the upper secondary group; *lit,* where the frequency was 55% for the native speakers, 35% for the university group, and 29% for the upper secondary group; and *twisted,* where the frequency was 41% for the native speaker group, 46% for the university group, and 25% for the upper secondary group. Some of the differences seen in cloze values for the upper secondary group compared to the native speakers and the university group can be linked to the upper secondary group not understanding the verbs. This is the case for *extinguished,* where ten participants reported that they did not know the word, and eight of these were blank responses. Despite some cloze values being affected by participants not understanding the verbs, there is a tendency for the cloze value to decrease in line with proficiency levels. The tendencies that have been reported so far will be further addressed in the discussion.

Table 4: Medium cloze probability verbs (66-34%), showing unique answer and cloze probability for each group.

Native speakers			University students			Upper secondary		
Medium	**Unique answer**	**Cloze**	**Medium**	**Unique answer**	**Cloze**	**Medium**	**Unique answer**	**Cloze**
Emptied	bin	0.66	Buckled	seatbelt	0.62	Brushed	teeth	0.63
Threw	ball	0.61	Ironed	shirt	0.62	Unfastened	seatbelt	0.63
Worshipped	God	0.61	Worshipped	God	0.62	Fried	chicken	0.58
Ironed	shirt	0.59	Ate	food	0.58	Threw	ball	0.54
Lit	candle	0.55	Kicked	ball	0.54	Ordered	pizza	0.50
Decorated	room	0.53	Arrested	thief	0.50	Juggled	ball	0.46
Fried	egg	0.53	Hunted	deer	0.50	Kicked	ball	0.46
Tied	shoelace	0.53	Knocked	(on the) door	0.50	Painted	house	0.46
Juggled	ball	0.51	Boiled	water	0.46	Unlocked	door	0.42
Scratched	head	0.51	Changed	clothes	0.46	Worshipped	God	0.42
Stroked	cat	0.50	Decorated	room	0.46	Boiled	egg	0.38
Knitted	scarf	0.44	Juggled	ball	0.46	Knocked	(on the) door	0.38
Owned	car	0.43	Licked	ice cream	0.46	Loaded	gun	0.38
Licked	ice cream	0.41	Threw	ball	0.46	Washed	car	0.38
Twisted	ankle	0.41	Twisted	ankle	0.46	Watched	game	0.21
Buckled	seatbelt	0.40	Collected	money	0.38			
Landed	plane	0.39	Cut	hair	0.38			
Typed	essay	0.39	Fried	chicken	0.38			
Fired	gun	0.36	Planted	tree	0.38			
Painted	wall	0.35	Loaded	gun	0.38			
Sipped	tea	0.34	Stroked	cat	0.38			
			Lit	fire/ candle	0.35			
			Married	him	0.35			
			Owned	car	0.35			
			Tied	shoelace	0.35			

Looking at the cloze probabilities of some of the verbs, one can see that there are cases where verbs have a very high cloze probability value, despite having selectional requirements that seem quite open, i.e. that are not very constraining. While the cloze probability for verbs like *extinguished, milked* and *hammered* reflect the constraint of their selectional requirements, the cloze values for verbs like *kicked, read,* and *emptied* do not. A corpus search was conducted for the verbs that displayed a high cloze probability in the native speaker group *(extinguished*, *milked, hammered, watered, kicked, unlocked, unfastened, read,* *knocked, brushed,* and *emptied),* to see if the frequency ranks in the corpus correlate with the cloze probability of the verbs and their most frequent filler.

Table 5: Frequency ranks from corpus search for verbs with a high cloze probability in the native speaker group.

extinguished_v NOUN	**141**	**extinguished_v * NOUN**	**1748**	**extinguished_v * * NOUN**	**1340**
extinguished_v fire	8	extinguished_v * fire	435	extinguished_v * * fire	81
extinguished_v fires	7	extinguished_v * fires	25	extinguished_v * * fires	12
	10.64 %		*26.32 %*		*6.94 %*
milked_v NOUN	**152**	**milked_v * NOUN**	**875**	**milked_v * * NOUN**	**675**
milked_v cow	4	milked_v * cow	40	milked_v * * cow	22
milked_v cows	71	milked_v * cows	114	milked_v * * cows	19
	49.34 %		*17.60 %*		*6.07 %*
hammered_v NOUN	**590**	**hammered_v * NOUN**	**2182**	**hammered_v * * NOUN**	**2075**
hammered_v nail	0	hammered_v * nail	34	hammered_v * nail	70
hammered_v nails	15	hammered_v * nails	22	hammered_v * * nails	19
	2.54 %		*2.57 %*		*4.29 %*
watered_v NOUN	**93**	**watered_v * NOUN**	**934**	**watered_v * * NOUN**	**1207**
watered_v plant	0	watered_v * plant	11	watered_v * * plant	8
watered_v plants	8	watered_v * plants	25	watered_v * * plants	8
	8.60 %		*3.85 %*		*1.33 %*
kicked_v NOUN	**1846**	**kicked_v * NOUN**	**12628**	**kicked_v * * NOUN**	**14257**
kicked_v ball	18	kicked_v * ball	1379	kicked_v * * ball	138
kicked_v balls	22	kicked_v * balls	29	kicked_v * * balls	42
	2.17 %		*11.15 %*		*1.26 %*
unlocked_v NOUN	**555**	**unlocked_v * NOUN**	**2024**	**unlocked_v * * NOUN**	**1438**
unlocked_v door	4	unlocked_v * door	221	unlocked_v * * door	81
unlocked_v doors	20	unlocked_v * doors	51	unlocked_v * * doors	18
	4.32 %		*13.44 %*		*6.88 %*
unfastened_v NOUN	**0**	**unfastened_v * NOUN**	**25**	**unfastened_v * * NOUN**	**29**
unfastened_v seat belt	0	unfastened_v * seat belt	5	unfastened_v * * seat belt	0
unfastened_v seat belts	0	unfastened_v * seat belts	0	unfastened_v * * seat belts	0
	0 %		*20.00 %*		*0 %*
read_v NOUN	**35603**	**read_v * NOUN**	**191775**	**read_v * * NOUN**	**122955**
read_v book	136	read_v * book	10778	read_v * * book	2847
read_v books	3335	read_v * books	4297	read_v * * books	2090
	9.75 %		*7.86 %*		*4.02 %*
knocked_v NOUN	**1118**	**knocked_v * NOUN**	**6472**	**knocked_v * * NOUN**	**10435**
knocked_v door	3	knocked_v * door	190	knocked_v * * door	3450
knocked_v doors	18	knocked_v * doors	527	knocked_v * * doors	459
	1.88 %		*11.08 %*		*37.46 %*
brushed_v NOUN	**730**	**brushed_v * NOUN**	**2942**	**brushed_v * * NOUN**	**2780**
brushed_v hair	8	brushed_v * hair	102	brushed_v * * hair	23
	0.82 %		*3.47 %*		*0.77 %*
brushed_v teeth	6	brushed_v * teeth	267	brushed_v * * teeth	2
	1.10 %		*9.08 %*		*0.07 %*
emptied_v NOUN	**249**	**emptied_v * NOUN**	**2217**	**emptied_v * * NOUN**	**1371**
emptied_v bin	0	emptied_v * bin	11	emptied_v * * bin	16
emptied_v bins	7	emptied_v * bins	17	emptied_v * * bins	8
	2.81 %		*1.26 %*		*1.75 %*

Investigations in the corpus showed that for most of the verbs, the NP listed as the most frequent in the corpus corresponded to the most frequent NP-fillers for the native speaker group. This was the case for *extinguished, milked, unlocked, unfastened, watered,* and *knocked.* For *brushed,* both *teeth* and *hair* were of the most frequent NPs, but *teeth* was more frequent than *hair.* 26% of the native speaker group responded with *teeth,* 31% of the university group responded with *teeth,* and 21% of the upper secondary group responded with *hair.* Hence, *teeth* was the second most frequent NP-filler for the native speakers and the university group, and *hair* was the second most frequent NP-filler for the upper secondary group. *Ball* was listed as the most frequent NP for *hammer* (for the search *hammered_v * NOUN),* however, in the search for two words between the verb and the noun (*hammered_v * * NOUN*), *nail* is of the most frequent NPs. The selectional requirements of these verbs can be considered as constrained, and this is reflected in both the cloze probability of the verbs and the frequency ranks in the corpus.

Despite the openness of the selectional requirements of *kick,* this verb had a high cloze probability. This was also reflected in the corpus, listing *ball* as the most frequent NP (for the search *kicked_v * NOUN).* For *read,* which also seems quite open in what it can select as its argument, *comment* was the most frequent NP, although *book* was of the most frequent NPs. Hence, the findings in the corpus can be thought to correspond with the frequency of the NP-fillers for *kick* and *read.* For *empty,* however, this was not the case. The corpus search listed *tank* as the most frequent NP for *empty* (*emptied_v * NOUN)*, and *bin* was listed far down on the list. Looking at table 5, one can see that both *kicked* and *read* are frequent verbs in the input. *Emptied* does not display this same level of frequency in the input.

The results and different tendencies that have been identified in this chapter will be further addressed in the following, exploring sensitivity to argument structure with regards to two different approaches.

5 Discussion

The aim of this thesis was to look at how sensitivity to argument structure is manifested and develops in Norwegian second language learners of English, and how the sensitivity of an L2 learner differs from that of a native speaker. First, it was important to establish that the two second language groups were in fact different with regards to proficiency. The mean values and standard deviation values of the grammar and vocabulary tests show that there is a significant difference between the groups in level of proficiency. The mean values for how many participants per verb were unsure of whether their completion of the sentence was grammatically correct, for how many that did not understand the sentence, and for how many that gave a blank response (table 2) were higher for the upper secondary group than for the university group across all three factors. This can be thought to reflect the differences that can be seen between the L2 groups in proficiency.

The STM (forwards digit recall) and the WM (backward digit recall) scores of the two groups of L2 learners show a significant difference on the outset. Thus, any results reflecting differences in sensitivity to second language argument structure may simply be attributed to the observed difference in working memory capacity, as language competence and language learning outcomes have been shown to correlate significantly with STM and WM capacity (Baddeley, Gathercole, & Papagno, 1998; Perani, 2005; Vulchanova et al., 2014). However, we believe that the sample for the current study is not very likely to be biased, since participants were invited to freely join the study, and were not specifically selected. Thus, the two groups reflect the natural variation that can be found in students at a lower L2 proficiency level and, respectively, those at a higher L2 proficiency level.

An alternative interpretation of the observed differences in WM scores between the two groups can, in fact, be in terms of the effect of (longer) language learning and exposure to the second language over time. It has indeed been shown that bilingualism and the acquisition of more languages can confer a working memory advantage (Morales, Calvo, & Bialystok, 2013). Thus, the university students who have had longer exposure to English, might have benefited from the process of learning English for a longer period of time, as reflected in better STM and WM capacity. Since this study did not set out to investigate this issue specifically, we cannot make any conclusions based on the current findings.

The performance on the incomplete sentences was expected to show a difference between the two L2 groups in that that the university group would reflect a similar level of constraint on verb argument structure to that of the native speaker group, and that the upper secondary group would show less sensitivity to argument structure. This expectation will be explored in the following with two different approaches – theoretical and computational.

5.1 Theoretical approach to sensitivity

As mentioned in the introduction, the first approach is theoretical, and is concerned with the idea that the selectional requirements of a verb are lexically encoded in the entry of the verb, and that this information is activated upon recognition of the verb and used to process and predict upcoming input in sentence processing (Koenig et al., 2003).

In line with the assumption made by Koenig et al. (2003), the incomplete sentences were expected to yield responses that would reflect the lexically encoded information about the verb and its argument structure, i.e. the selectional requirements set by the verb. This was intended to serve as an indicative of sensitivity to the verbs argument structure. The most frequent unique answers of both L2 groups display a general tendency to meet the requirements set by the verbs, thus, indicating a sensitivity to argument structure. However, the cloze probabilities for the two groups, compared to the native speaker group, could indicate a difference. Level of constraint can be seen as a property of the verb, which is caused by the selectional requirements or argument structure set by the verb. Hence, high constraint verbs should in turn display high cloze probabilities. This is reflected in the cloze values for the native speaker group. Looking at the verbs listed in the high constraint category (table 3), one can see through the cloze values that the university group display a similar constraint on the verbs argument structure as the native speaker group. The university group only lack two verbs in this category, but the cloze values for *kicked* and *knocked* are still quite high in the medium cloze category. For the upper secondary group, however, this is not the case. The upper secondary group's cloze values for the high constraint verbs (as reflected by the native speaker group) are lower compared to both the native speakers and the university students. This difference between the two L2 groups, and between the upper secondary group and the native speaker group, could indicate that the university students are somewhat similar to the native speakers in sensitivity to the lexical specifications of the verbs, and that the upper secondary group, in turn, are not as sensitive to this as the native speakers and the university group.

The tendency of the university group to display a constraint on the verb argument structure similar to that of the native speakers, and the general tendency of the upper secondary group to show a wider distribution, may be compatible with the combined results from the studies conducted by Reine (2016) and Johnsen (2016). The results from Reine (2016) show a higher mean of proportions of gazes towards the target item in the most constrained verb category (for the target present condition). Looking at the same results for Johnsen (2016) shows that proportion of looks towards the target item in the upper secondary group was also higher for the most constrained verbs, but a comparison shows that the mean value was lower for the upper secondary group than for the university group. This may be compatible with the constraint results for this study, where the university group is closer to the native speaker group, and the upper secondary group shows a lower level of constraint on the verbs argument structure. This comparison can only be seen as implications, since a statistical comparison has yet to be made between the results of Johnsen (2016) and Reine (2016).

More differences between the two L2 groups can be found upon closer inspection of the unique responses for some of the verbs, which reveals certain tendencies of some participants not meeting the selectional requirements set by the verb. This was mostly the case for the upper secondary group, suggesting that the sensitivity is indeed somewhat different for this group compared to the university group. These cases of responses not meeting the selectional requirements can be interpreted as signs of transfer from L1 representations to L2.

Based on the arguments made by Stringer (2010) regarding lexical relativity and the challenges that an L2 learner faces when learning a new language, it was expected that the upper secondary group would show signs of transfer from their L1 lexicon. There are some cases which indicate that the upper secondary group, and sometimes the university group as well, treat the verbs differently from the native speaker group. While the sizes of the two L2 groups are not large enough to make any generalizations on lexical transfer, there are some verbs which show tendencies of transfer of L1 representations to L2. The following sections will take a closer look at these verbs, with the intention to try and understand why some participants show a different treatment of the verbs. (An overview for all the unique responses for the verbs that will be discussed can be found in Appendix C).

5.1.1 Itched vs. scratched

The university and upper secondary group responses for the incomplete sentence *"He itched…"* were to a certain degree similar to the responses for *"Ronald scratched"*. The most frequent NP-filler for *"Ronald scratched…"* was *back* for both the university group (15%) and the upper secondary group (17%). *Back* was also the most frequent NP-filler for *"He* *itched…"* in the upper secondary group (29%), even though the verb cannot take this NP-filler as its argument. While the majority of the university group participants treated *itch* either as an intransitive verb, or as a verb taking an infinitive clause, several of the responses indicated a similar treatment of *itch* and *scratch,* with NP-fillers such as *neck, back, shoulders,* etc. Overall, the majority of the university group responses were adjuncts (PPs) or infinitive clauses; the majority of the upper secondary responses were NP-fillers treated as arguments.

The difference between the two English verbs is clear. *Scratch* is a transitive verb, denoting an action/event, and as stated in the description of the selectional requirements of the verb, it must take either a location argument, selecting an animal or human body part to fill the slot, or a patient argument, selecting almost any entity (in the sense of scratching something and leaving a mark, like scratching a car). *Itch*, however, can in this case be an intransitive verb denoting a state, not taking an argument in the object slot, or it can denote a desire to do something, a more informal sense of the verb, taking an infinitive clause *("He itched to do something").* While the tendency was stronger in the upper secondary group, the responses from both groups imply a treatment of *itch* as similar to *scratch,* where both are given a location argument in the object slot. The link that can be seen in the upper secondary group is most likely due to their lexical knowledge of the Norwegian verb *"å klø",* which is analogous for both *itch* and *scratch*. This can be exemplified through the following examples:

(8) "Ronald **klødde** seg på nesen".
Ronald scratched his nose.

(9) "Nesen hans **klødde"**.
His nose itched.

"Å klø" can be both intransitive and transitive, denoting the state of something itching, or the event where an agent scratches someone or something (location/theme). This distinction between the transitive and intransitive, event and state, is denoted in two separate verbs in English, *scratch* and *itch*. This treatment of *itch* might serve as evidence for lexical transfer from L1 representations to L2, as discussed by Stringer (2010). The responses for both the university group and the upper secondary group imply a tendency to overgeneralize the use of *itch,* applying the selectional restriction of *scratch* to *itch,* based on the Norwegian verb *"å klø".* Comparing the two L2 groups show a stronger tendency of lexical transfer in the upper secondary group than in the university group.

5.1.2 Loaded

Another example which indicates transfer from L1 to L2 is the verb *load,* in the incomplete sentence *"Mary loaded...".* The tendency is not as clear in this case as with *itch/scratch,* but it might still imply an overgeneralization or transfer. In the upper secondary group, two participants completed the sentence with the NP-filler *telephone,* and one responded with *batteries.* "*To load"* can be translated both into *"å laste"* and *"å lade"* (which again can be translated into *"to charge"*). While the Norwegian verb *"å lade"* can denote the event of filling something up with electricity (like the English verb *"to charge"),* the English verb *"to load"* does not. There is, however, a crossing between *"to load"* and *"å lade"*, which can be illustrated in the following examples:

(10) "Han **ladet** pistolen".
He loaded the gun.

(11) "Han **ladet** telefonen sin".
He charged his telephone.

While English uses two different verbs to denote two different events, the Norwegian verb *"å lade"* uses the same verb to denote two different events (different senses). This difference can imply that some participants have treated *"to load"* and *"å lade"* the same way, with two senses, taking two different arguments. This might indicate a lexical transfer between the L1 lexical representation and the L2. Comparing the most frequent NP-filler for the three different groups can support this theory. While the native speaker group had *car* (30%) as their most frequent unique answer, both the university group and the upper secondary group had *gun* (both 38%) as their most frequent answer. 15% responded with *gun* in the native speaker group, 23% responded with *car* in the university group, and 12% responded with *car* in the upper secondary group. This difference in NP-fillers between the native speakers and the L2 learners might be due to exposure to input in L2. It might also imply a stronger activation of the sense of the verb which is linked to the Norwegian verb *"å lade",* which can justify the transfer.

5.1.3 Served

The unique answers for the sentence *"John served..."* in the upper secondary group show a slight different use from those of the native speakers. Some participants completed the sentence with a PP, as illustrated in the following sentences:

(12) John served in a wedding.

(13) John served at a party.

(14) John served at a restaurant.

Not a single one of the unique answers from the native speaker group show this type of response, indicating that some participants in the upper secondary group treat the verb the same way as the Norwegian verb *"å servere"*. As described earlier, in the selectional requirements of the verb, the English verb *"to serve"* can select as an argument a human being or group of human beings, an organization (providing a service for someone), or something edible (both solid and liquid). In Norwegian, *"to serve"* can be translated to both *"å servere"* and *"å tjene"*. *"Å servere"* means to serve or bring food to someone, and can be realized as an intransitive verb. While *"to serve"* also can be realized as an intransitive verb, as in *"He served at the table",* the responses from the native speaker group imply that this is not a frequent realization of the verb.

Hence, the tendency of some upper secondary participants to treat the verb *"to serve"* the same way as *"å servere"* may be due to transfer. Statistical learning might be an explaining factor for the difference between the university group and the upper secondary group, where more exposure to relevant input may have allowed the university participants to develop a native like understanding of the verb, adapting the treatment and use of the verb to that of a native speaker.

5.1.4 Odd cases

There are some tendencies of the upper secondary group understanding verbs as something different from what they really are, simply because they look like Norwegian verbs, or are similar to another English or Norwegian word (possibly another word class). The verb *examine* showed such tendencies, exemplified in (15)-(18):

(15) I examined in Spanish.

(16) I examined in math.

(17) I examined last year (the response for three participants).

(18) I examined today.

These responses to the verb *examine* indicate an intransitive treatment of the verb. The treatment of the verbs discussed above (*itch, load,* and serve) could be explained by looking at analogous verbs in Norwegian. For *examine,* however, this is not the case. What might have happened here is that some upper secondary participants have linked the verb *examine* and the noun *exam,* thinking that *examine* means to take an exam. The most reasonable translation for the examples above are given in the following:

(19) ”Jeg tok eksamen i spansk”.
I had an exam in Spanish.

(20) ”Jeg tok eksamen i matte”.
I had an exam in math.

(21) ”Jeg tok eksamen i fjor”.
I took the exam last year.

(22) “Jeg tok eksamen i dag”.
I took the exam yesterday.

The same type of examples can also be found for the incomplete sentence *“James fired…”.* Some of the participants responded with NPs such as *“the pictures”, “a hotdog”,* or *“the house”,* which may be explained by the link between the verb and the noun *“to fire”* and *“a fire”.* These are only assumptions, and taking a closer look at the responses shows that participants giving these responses also reported that they were unsure of their answer. This was especially the case for *examine.* Hence, it might just be the case that they did not understand the meaning of the verb, and simply made a guess based on a noun which was familiary to them. The performance on the incomplete sentences has in fact shown that several of the verbs were not part of the vocabulary for many of the upper secondary participants, meaning that they never recognized the verb in the first place.

The cases of transfer seen with the verbs discussed above are only tendencies, and no conclusions can be drawn from them. However, it is possible to consider them as implications of features that can be seen in the L2 language acquisition process. All the verbs discussed above – *itch/scratch, load* and *serve* – show the tendency of some participants to treat the English L2 verbs the same way as the analogous Norwegian verbs, by treating the L2 target verbs as having the same argument structure as the Norwegian analogue. This can imply that the lexical representations of their L1 also serve as representation for their L2 items, or as Stringer (2010) argues, the L1 representations serve as the initial state for the L2 items. The differences that can be seen between the university group and the upper secondary group with regards to lexical transfer may be due to exposure to input and statistical learning. As pointed out by Treffers-Daller and Calude (2015), statistical learning is an important tool for L2 learners in overcoming and retreating from overgeneralizations and transfer from L1 to L2. Hence, more exposure to input and knowledge of the frequency of words and the patterns they appear in may have aided the university group in adjusting their use of the verbs to a level which is similar to the native speakers.

To summarize, the performance on the incomplete sentences, with regards to meeting the selectional requirements set by the verbs, did not yield a large difference between the university group and the upper secondary group, indicating that both L2 groups are sensitive to the argument structure of the verbs. However, the distribution of verbs according to level of constraint, and the cloze probability values presented for each verb in all three groups suggest that the upper secondary group are less sensitive to the constraint of the argument structure of the verbs, and that the university group have more or less the same constraint on argument structure as the native speakers. Furthermore, as was expected, the upper secondary group did show more signs of transfer. Statistical information and exposure to input for a longer period of time may have allowed the university group, and many of the upper secondary participants, to retreat from overgeneralization and change their L2 representations. In terms of sensitivity to argument structure in sentence processing and transfer, this would mean that upon recognition of the verb, the information that is activated will be based on the L1 representation of the verb, affecting the prediction of the upcoming input.

5.2 Computational approach to sensitivity

The computational approach assumes that arguments are represented at a more concrete level, whereby native speakers over time form associations between a specific verb and the possible fillers for the NP slot, and would typically activate NP-fillers which are most frequently used. We know from research that frequency is an important feature in language acquisition, and that humans are aware of the frequency of the words in our language (Ellis, 2002). We also know that this type of knowledge about frequency can be found about second language in second language learners (Schmitt & Dunham, 1999). Further, different studies have shown that there is a frequency effect in word recognition, where words that are more frequent in the input have stronger activations in sentence processing and word recognition (Forster and Chambers, 1973, and Marslen-Wilson, 1987, in Altmann, 1998). The cloze probabilities displayed in this study could imply a frequency effect in what is activated upon recognition of the verb.

As described in the results, there is a tendency for the cloze probability values to decrease with level of proficiency. This tendency could imply a frequency effect, where the university group display a frequency of the NP-fillers for the verbs which is similar or close to that of the native speakers, and the upper secondary group have not yet developed this same knowledge about frequency and what is a typical NP for the argument slot. These indications can be supported by the findings of Schmitt and Dunham (1999), where it was found that moderately advanced second language learners in fact can acquire a knowledge about the frequency of words similar to native speakers.

On closer inspection of the cloze values and the most frequent NP-fillers, there are features of the results which can strengthen the idea of a sensitivity to frequency. As stated earlier, the cloze values for the native speaker group can be thought to reflect the level of constraint in the verbs argument structure. This is the case for most of the verbs with a high cloze probability in the native speaker group. However, in some cases, verbs that in theory have very open selectional requirements, i.e. does not have a high level of constraint, actually convey a high cloze probability. This will be discussed in the following.

A comparison of the verbs with a high cloze probability, and the descriptions of the selectional requirements of the verbs can provide some interesting information as to what is actually activated upon recognition of the verb. Verbs like *milk* appear to be very restricted:

(23) *Laura milked* – [PATIENT] argument, most likely to select an animal (mammal) that produces milk.

This is reflected in the high cloze probability for *milk* across all three groups (native speakers, 93%; university group, 92%; upper secondary group 88%). Also, there are very few unique answers in total for each group (native speakers, 5; university group, 3; upper secondary, 4), reflecting the narrow selectional requirements of *milk.*

In comparison, verbs like *kick* are not as restrictive in their selectional requirements:

(24) *She kicked* – [THEME/LOCATION] argument (Levin, 1993; Saeed, 2009), can select almost any physical entity, human or non-human, that can be kicked using your leg and foot.

Despite this, the responses of the native speaker group have placed *kick* in the high cloze category, with the most frequent unique answer being *ball* (75%, 16 unique answers out of 101 responses). This noun matches the criteria set by the verb's argument structure, but so would almost anything else – *kicked the wall, kicked her brother, kicked the door, kicked the bucket*. Neither of the two L2 groups show this same restrictedness. Both groups placed *kicked* in the medium cloze category, where 54% of the university group responses were *ball* (10 unique answers out of 26 responses), 46% for the upper secondary group (12 unique responses out of 24 responses). Still, the tendency of both the native speakers and the two L2 groups to use *ball* as frequently as they do, despite the open selectional requirements of the verb, could indicate that frequency is an important factor in what is actually activated upon recognition if the verb.

Read and *empty* can also be considered as having quite open selectional requirements:

(25) *She read* – [THEME] argument, can select an argument that is any form of literature or written text.

(26) *He emptied* – [PATIENT/LOCATION] argument (Levin, 1993), an entity that is a container of some sort, that can contain something (*the trash, his mind*, etc.).

Both of these verbs appear open in selectional requirements, but the cloze values for the native speaker group imply that the verbs have a constrained argument structure. The tendency for the verbs to display such high cloze values, despite the openness of their selectional requirements, imply that there must be some frequency effect in play. The corpus search that was conducted can somewhat justify this suspicion. As summarized in the results, the corpus search showed that for the high cloze verbs that appear restricted in selectional requirements, the most frequent NP-filler for the native speaker group match the findings in the corpus, e.g. *fire* was listed as the most frequent NP to serve as an object for *extinguished.*

This correlation between native speaker responses and the corpus could also be found for some of the open verbs with a high cloze value. The search in the corpus showed that *ball* was the most frequent NP for *kicked,* and *book* was of the most frequent NPs for *read.* Hence, the high frequency of occurrence in the language for the verb *kicked* with *ball,* and *read* with *book*, as an NP in the argument slot appears to be reflected in performance on the incomplete sentences. The corpus did, however, not show this correlation for *emptied* and *bin.* This may be due to the type of corpus that was used, and that *"He emptied the bin"* might not be a sentence which is frequently used in news on the web. This was reflected when comparing the ranks in table 5. It is likely that a phrase like *"He emptied the bin"* is more frequent in spoken language, rather than written, and will therefore not appear frequent in the corpus. This could also justify why the two L2 groups displayed such different cloze values from the native speaker group. The university group had a cloze value of 23% for *emptied*, with *trash* being the most frequent unique answer. The upper secondary group had a cloze value of 17%, where *bottle/pocket* were the most frequent unique answers.

What does this mean for sensitivity to argument structure? Based on what has been addressed so far, the results can possibly suggest that the native speakers have formed an association between the specific verbs and the possible fillers for the NP slot, and will therefore typically activate NP-fillers which are more frequently used. This could mean a more concrete representation of the arguments of verbs in their entries. If it is in fact the case that native speakers over time can form this association between some verbs and a possible filler, the same could possibly be said for the two L2 groups, especially for *read* and *ball.* However, this may be due to frequency in the input of the L1. Regardless, this is an interesting implication, and should be further investigated in future research.

There are several factors that could have caused the differences in frequency of the NP-fillers. As addressed earlier in the discussion, the participants for the upper secondary group appeared to be more insecure of their responses, and some verbs' cloze values have been affected by the fact that many participants did not understand the verb. The frequency of the NP-fillers, and thus, also the cloze values, could also have been affected by factors such as frequency of occurrence in L1 input (Norwegian), frequency in L2 input (English), and the amount of exposure to input. It could also be affected by cultural factors, such as social environment and differences between generations.

To summarize the findings in this approach, there is a general tendency for the cloze values of the verbs to decrease with level of proficiency. This indicates a sensitivity to frequency which is more evident in the university group than the upper secondary group. The high cloze values for some verbs which appear open in selectional requirements could indicate that the native speakers have formed an association between the specific verb and the NPs which are most frequent in the language. This can suggest that when we recognize the verb, we might also activate information about the most frequent NP to occur as the argument for the verb. While the results could imply such a sensitivity, it cannot be concluded, and more research on the topic would be required.

5.3 Nature of the method

The methodological design of the thesis has both strengths and limitations. An important strength of using incomplete sentences is that it offers a considerable amount of material, which can be used to look at several features regarding argument structure (e.g. sensitivity, frequency of NP-fillers, transfer). However, having the participant respond to incomplete sentences is a behavioral method, which means that we cannot control how they interpret the sentence and the verb. Furthermore, the material was of such nature that we were not able to do a quantitative analysis of correlations between the proficiency level and working memory capacity of the participants, and their performance on the incomplete sentences. Therefore, we cannot make any generalizations on the results regarding sensitivity and proficiency levels. One can, however, see indications and make inferences in the larger differences.

While the incomplete sentences do provide a great amount of material that can be analyzed, the two L2 groups are very much smaller in size compared to the native speaker group, which makes it difficult to make an exact comparison. For example, it is quite random what is the most frequent response in the low cloze probability verbs, considering the small size of the L2 groups. Also, one L2 participant's response to an incomplete sentence has more of an effect on the cloze probability value than what a native speaker's response would have. Therefore, the differences that can be seen in the cloze value across the groups cannot be considered as exact measures.

6 Conclusion

This thesis has explored two different approaches to the development of sensitivity to argument structure in second language learners. The expectations to the study were met in that there seems to be a difference in sensitivity to argument structure between the two L2 proficiency groups. The findings made throughout the study could indicate that as proficiency level increases, so does the sensitivity to the constraint of the verbs. The cloze probability of the verbs imply that the upper secondary group are less sensitive to the constraint of the selectional requirements than the university group. The university group show a constraint on argument structure which is close or similar to the native speaker group. Furthermore, while both L2 groups display a sensitivity to argument structure in that they meet the selectional requirements of the verbs, the upper secondary group show more signs of transfer from L1 representations to L2. Concerning the computational approach, it has been suggested that the university group are more sensitive to the frequency of typical NP-fillers than the upper secondary group. The cloze values for some of the verbs, which appear open in selectional requirements, could suggest that the native speakers have formed some kind of association between the specific verb and the most typical NP.

The methodological design of this thesis has provided a considerable amount of material that can be qualitatively analyzed, and contributes with insight into sensitivity to argument structure in second language learners. Hence, it opens for several research questions that should be further examined. Since the scope of the thesis has been somewhat exploratory, more research should be conducted to provide evidence for some of the indications and suggestions that have been made. The concept of sensitivity to argument structure in second language learners should be further addressed, as there is little research similar to that which has been conducted on native speakers.

References

Aitchison, J. (2003). *Words in the mind : an introduction to the mental lexicon* (3rd ed. ed.). Oxford: Blackwell.

Altmann, G. T. (1998). *The ascent of Babel: an exploration of language, mind, and understanding*: OUP Oxford.

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247-264.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological review, 105*(1), 158.

Block, C. K., & Baldwin, C. L. (2010). Cloze probability and completion norms for 498 sentences: Behavioral and neural validation using event-related potentials. *Behavior Research Methods, 42*(3), 665-670. doi:10.3758/brm.42.3.665

Bloomfield, L. (1933). *Language*. New York: Holt.

Boland, J. E. (2005). Visual arguments. *Cognition, 95*(3), 237-274. doi:http://dx.doi.org/10.1016/j.cognition.2004.01.008

Dowty, D. (1991). Thematic proto-roles and argument selection. *language, 67*(3), 547-619.

Ellis, N. C. (2002). Frequency effects in language processing. *Studies in second language acquisition, 24*(02), 143-188.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of verbal learning and verbal behavior, 12*(6), 627-635.

Friederici, A. D., & Frisch, S. (2000). Verb argument structure processing: The role of verb-specific and argument-specific information. *Journal of Memory and language, 43*(3), 476-507.

Garrod, S. (2006). Psycholinguistics: Psyycholinguistic Research Methods. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (Second ed.). Amsterdam: Elsevier.

Haegeman, L. (1994). *Introduction to government and binding theory* (2nd ed. ed. Vol. 1). Oxford: Blackwell.

Haegeman, L., & Guéron, J. (1999). *English grammar: A generative perspective*: Blackwell Publishing.

Jackendoff, R. (2002). What's in the Lexicon? *Storage and computation in the language faculty*, 23-58.

Johnsen, R. (2016). *Sensitivity to second language argument structure: an experimental study with Norwegian learners of English.* (Master), NTNU, Trondheim.

Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and language, 49*(1), 133-156.

Koenig, J.-P., Mauner, G., & Bienvenue, B. (2003). Arguments for Adjuncts. *Cognition: International Journal of Cognitive Science, 89*(2), 67-103. doi:10.1016/S0010-0277(03)00082-9

Levin, B. (1993). *English verb classes and alternations : a preliminary investigation*. Chicago: University of Chicago Press.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*(1), 71-102.

Meints, K., Plunkett, K., & Harris, P. L. (2008). Eating apples and houseplants: Typicality constraints on thematic roles in early verb learning. *Language and Cognitive Processes, 23*(3), 434-463.

Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of experimental child psychology, 114*(2), 187-202.

Perani, D. (2005). The neural basis of language talent in bilinguals. *Trends in Cognitive Sciences, 9*(5), 211-213. doi:10.1016/j.tics.2005.03.001

Pickering, S., & Gathercole, S. E. (2001). *Working memory test battery for children - Manual*. London: Pearson Education Ltd.

Pinker, S. (2013). *Learnability and Cognition: The Acquisition of Argument Structure* (New edition. ed.): United States: Mit Press.

Reine, A. S. (2016). *L2 speakers' argument structure sensitivity inferring from the participant's eye gazes: an eye-tracking experimental study with adult Norwegian learners of English.* (Master), NTNU, Trondheim.

Saeed, J. I. (2009). *Semantics* (3rd ed. ed. Vol. 2). Chichester: Wiley-Blackwell.

Schmitt, N., & Dunham, B. (1999). Exploring native and non-native intuitions of word frequency. *Second Language Research, 15*(4), 389-411.

Stringer, D. (2010). The gloss trap. In Z. Han & T. Cadierno (Eds.), *Linguistic relativity in SLA: Thinking for speaking*. Bristol: Multilinguial Matters.

Treffers-Daller, J., & Calude, A. (2015). The Role of Statistical Learning in the Acquisition of Motion Event Construal in a Second Language. *International Journal of Bilingual Education and Bilingualism, 18*(5), 602-623. doi:10.1080/13670050.2015.1027146

Vulchanova, M., Foyn, C. H., Nilsen, R. A., & Sigmundsson, H. (2014). Links between phonological memory, first language competence and second language competence in 10-year‐old children. *Learning and Individual Differences, 35*, 87-95.

Appendix A – Low cloze probability verbs (33%-0%)

Native speakers			University students			Upper secondary		
Low	**Unique answer**	**Cloze**	**Low**	**Unique answer**	**Cloze**	**Low**	**Unique answer**	**Cloze**
Arrested	criminal	0.33	Fired	gun	0.31	Arrested	thief	0.33
Tied	shoelace	0.32	Knitted	scarf	0.31	Decorated	room	0.33
Boiled	kettle	0.31	Sipped	tea	0.31	Hunted	deer	0.33
Mended	bicycle	0.31	Annoyed	brother	0.27	Ironed	shirt	0.33
Moved	house	0.31	Frightened	him	0.27	Licked	ice cream	0.33
Ate	food	0.30	Painted	picture/ house	0.27	Missed	goal	0.33
Collected	stamps	0.30	Pinched	cheek	0.27	Watered	plant	0.33
Loaded	car	0.30	Pushed	him	0.27	Carried	bag	0.29
Squeezed	lemon	0.30	Squeezed	lemon	0.27	Cut	finger	0.29
Tightened	seatbelt	0.30	Borrowed	book	0.23	Delivered	mail	0.29
Pushed	door	0.29	Delivered	mail	0.23	Extinguishe d	fire	0.29
Visited	grandmother	0.29	Emptied	trash	0.23	Itched	back	0.29
Delivered	parcel	0.28	Guarded	door	0.23	Lit	fire	0.29
Hunted	fox	0.28	Landed	plane	0.23	Planted	flowers	0.29
Ordered	pizza	0.28	Obeyed	law	0.23	Pushed	him	0.29
Planted	seed	0.28	Played	game	0.23	Typed	message	0.29
Washed	dishes	0.25	Tightened	grip	0.23	Annoyed	brother	0.25
Borrowed	book	0.24	Wore	sweater	0.23	Broke	phone	0.25
Carried	bag	0.24	Broke	window	0.19	Changed	hair	0.25
Guarded	door	0.24	Carried	bags	0.19	Served	food	0.25
Missed	bus	0.24	Filled	cup	0.19	Squeezed	ball	0.25
Watched	football	0.24	Impressed	teacher	0.19	Tied	shoelace	0.25
Cut	hair	0.23	Memorized	word	0.19	Twisted	ankle	0.25
Entertained	friend	0.23	Moved	away	0.19	Borrowed	sweater	0.21
Dropped	phone	0.22		to the city	0.12	Buried	dog	0.21
Served	customer	0.22		car	0.08	Dropped	phone	0.21
Changed	clothes	0.21	Ordered	him	0.19	Filled	cup	0.21
Pinched	himself	0.21		pizza	0.15	Knitted	sweater	0.21
Reduced	price	0.21	Visited	parents	0.19	Played	football	0.21
Impressed	boss	0.20	Washed	hair/ clothes	0.19	Bought	car	0.17
Ground	pepper	0.18	Bought	dress	0.15	Buckled	seatbelt/ up	0.17
Measured	height	0.18	Buried	body/ treasure	0.15	Emptied	bottle/ pocket	0.17
Played	piano	0.18	Entertained	guests	0.15	Entertained	mother	0.17
Memorized	words	0.17	Itched	all over	0.15	Frightened	child	0.17
Annoyed	brother	0.16	Judged	everyone	0.15	Impressed	teacher	0.17
Frightened	cat	0.16	Measured	distance	0.15	Married	friend/ man	0.17
Obeyed	law	0.16	Mended	wound	0.15	Memorized	words	0.17
Rubbed	head	0.16	Missed	target/ mother	0.15	Owned	car	0.17
Buried	body	0.15	Reduced	sentence	0.15	Pinched	arm	0.17
Examined	paper	0.15	Removed	clothes	0.15	Rubbed	back	0.17

Removed	clothes	0.14	Scratched	back/ neck	0.15	Scratched	back	0.17
Judged	competition/ friend	0.14	Typed	letter	0.15	Sipped	wine	0.17
Married	man	0.14		fast	0.27	Tightened	rope	0.17
Wore	shirt	0.12	Watched	game	0.15	Visited	grandmoth er	0.17
Approved	proposal	0.11	Approved	application /proposal	0.12	Approved	it	0.13
Broke	leg	0.10	Dropped	phone	0.12	Collected	stamp/ money	0.13
Filled	glass/ car	0.10	Examined	scene/ paper	0.12	Examined	last year	0.13
Bought	dress	0.09	Ground	beans/ beef	0.12		body	0.08
			Served	food/ sentence/ dinner	0.12	Fired	employee/ gun	0.13
			Rubbed	shoulder/ arm/ neck	0.08	Guarded	door/ house/ castle	0.13
						Hammered	nail/ plank	0.13
						Judged	boy/ her	0.13
						Landed	plane/ (on his) face/ ..feet)	0.13
						Obeyed	law	0.13
						Stroked	cat	0.13
						Wore	tie	0.13
						Measured	table	0.08
						Moved	to the city/ away	0.08
							furniture	0.04
						Reduced	speed/ income	0.08
						Removed	teeth/ tattoo/ arm	0.08
						Ground	daughter/ breaker/ child/ almonds/ company	0.04
						Mended	himself/ fence/ feelings/ cut	0.04

Appendix B – Mean values for each verb

Verb	Total u.a., native speakers	Total u.a., university	Unsure	Do not know	BLANK	Total u.a., upper secondary	Unsure	Do not know	BLANK
Annoyed	26	11	0	0	0	9	0	0	0
Approved	49	16	2	1	0	14	8	6	6
Arrested	25	7	0	0	0	12	3	0	0
Ate	21	9	0	0	0	9	3	0	0
Boiled	10	5	0	0	0	8	3	4	4
Borrowed	32	9	0	0	0	14	3	1	1
Bought	52	18	0	0	0	19	3	0	0
Broke	43	15	0	0	0	13	1	0	0
Brushed	8	3	0	0	0	6	1	0	0
Buckled	16	6	5	1	0	6	5	13	12
Buried	42	17	0	0	0	12	5	2	3
Carried	41	17	1	0	0	17	3	0	0
Changed	34	11	0	0	0	15	5	0	0
Collected	39	11	0	0	0	18	1	0	1
Cut	34	11	0	0	0	9	2	0	0
Decorated	12	11	0	0	0	6	0	0	0
Delivered	22	10	0	0	0	8	1	1	1
Dropped	30	20	1	0	0	17	1	0	0
Emptied	23	13	0	0	0	13	1	4	4
Entertained	22	18	0	0	0	13	1	0	0
Examined	60	19	1	0	0	17	5	4	4
Extinguished	4	5	5	0	0	10	4	10	8
Filled	37	14	1	0	0	16	1	1	1
Fired	25	14	0	0	0	17	5	0	1
Fried	13	10	2	0	0	9	2	1	1
Frightened	25	14	0	0	0	11	6	5	5
Ground	35	13	5	9	7	5	9	13	19
Guarded	40	15	2	0	0	14	4	4	4
Hammered	11	5	6	0	0	17	7	1	1
Hunted	34	12	3	0	0	15	3	0	0
Impressed	33	18	1	0	0	15	3	0	0
Ironed	13	6	0	0	0	7	5	7	10
Itched		20	9	0	1	13	3	4	4
Judged	46	19	2	0	0	16	4	2	2
Juggled	35	13	5	0	0	9	6	4	5
Kicked	16	10	0	0	0	12	0	0	0
Knitted	17	9	0	0	0	9	5	8	9
Knocked	19	7	1	0	0	8	3	0	0
Landed	24	14	1	0	0	17	1	0	0
Licked	19	10	1	0	0	13	3	2	2
Lit	9	9	0	1	0	6	3	9	8
Loaded	17	11	2	0	0	9	4	0	0
Married	27	12	0	0	0	17	1	0	0

Measured	44	18	0	0	0	16	4	7	7
Memorized	33	17	1	0	0	18	2	0	0
Mended	38	20	2	1	0	5	2	19	19
Milked	5	3	1	0	0	4	3	0	0
Missed	34	16	0	0	0	10	1	1	1
Moved	34	17	1	0	0	22	1	0	0
Obeyed	25	15	0	0	0	12	4	9	10
Ordered	33	16	2	0	0	11	4	0	0
Owned	24	11	0	0	0	15	3	0	0
Painted	25	10	0	0	0	8	3	0	0
Pinched	22	12	0	0	0	11	5	10	10
Planted	18	8	0	0	0	10	2	0	0
Played	37	16	0	0	0	17	2	0	0
Pushed	31	14	0	0	0	14	1	0	0
Read	25	7	0	0	0	7	6	0	0
Reduced	41	17	4	0	1	19	6	2	2
Removed	49	21	1	0	0	21	3	0	0
Rubbed	33	20	2	0	0	14	4	4	4
Scratched	21	15	0	0	0	12	6	3	5
Served	33	16	0	0	0	16	4	0	0
Sipped	19	9	2	0	0	14	5	3	3
Squeezed	21	15	0	0	0	12	3	0	0
Stroked	14	9	2	1	1	13	8	6	7
Threw	23	9	1	0	0	10	4	0	0
Tied	10	9	1	0	0	15	3	1	1
Tightened	23	11	1	0	0	12	3	6	6
Twisted	24	10	3	0	0	16	6	0	0
Typed	26	15	0	0	0	12	0	0	0
Unfastened	13	5	1	0	0	2	1	7	8
Unlocked	13	4	0	0	0	12	0	0	0
Visited	28	11	0	0	0	13	0	0	0
Washed	20	11	0	0	0	8	1	0	1
Watched	42	18	1	0	0	15	2	1	1
Watered	9	5	0	0	0	9	3	5	5
Wore	35	15	0	0	0	17	0	3	3
Worshipped	25	9	0	0	0	6	2	6	6

Appendix C – Total list of unique responses for verbs addressed in the discussion

Served (n.s.)	Number of responses	Served (uni.)	Number of responses	Served (u.s.)	Number of responses
customer	22	food	3	food	6
dinner	19	sentence	3	in a wedding	2
food	15	dinner	3	(in the) military	2
meal	4	tea	2	soup	2
tables	4	customers	2	turkey	1
client	3	guests	2	at a party	1
drink	3	in the	2	lobster	1
main	2	pie	1	pizza	1
guests	2	beer	1	KFC	1
time	2	to help his nation	1	customers	1
in the	2	main	1	dinner	1
lunch	2	time in jail	1	at the restaurant	1
beer	1	ice cream	1	meal	1
glass of wine	1	friends	1	people	1
life	1	family	1	president	1
stew	1	king	1	army	1
breakfast	1				
caviar	1				
chips	1				
for the	1				
friends	1				
signature dish	1				
people	1				
pork	1				
snails	1				
soup	1				
spaghetti	1				
the army	1				
ball	1				
man	1				
pints	1				
coffee	1				
tea	1				

Itched (uni.)	**Number of responses**	**Itched (u.s.)**	**Number of responses**
all over	4	back	7
back	2	BLANK	4
arm	2	ear	2
to do something	1	knee	1
to touch his nose	1	scratch	1
a lot	1	hand	1
to switch seats	1	on his arm	1
closer	1	on her pimple	1
BLANK	1	head	1
like crazy	1	all day	1
scab	1	stitch	1
pimple	1	badly	1
for work	1	himself	1
on the	1	shoulder	1
to begin	1		
on the	1		
neck	1		
scratch	1		
groin	1		
on his elbow	1		
shoulder	1		

Loaded (n.s.)	Number of responses	Loaded (uni.)	Number of responses	Loaded (u.s.)	Number of responses
car	30	gun	10	gun	9
gun	16	car	6	truck	5
van	13	truck	3	car	3
dishwash	11	cart	1	telephone	2
washing machine	10	boxes into the car	1	ship	1
truck	5	cargo	1	batteries	1
boot	4	off her burden	1	movie	1
car boot	2	up on	1	trunk	1
lorry	2	trunk	1	something	1
box	1	PC	1		
computer	1				
dishes	1				
game	1				
shelves	1				
the shopping in her car	1				
trolley	1				
tumble dryer	1				

Appendix D – Incomplete sentences

Part one

Information

Thank you for participating! First we need some basic information about you. Please fill in the following, and then proceed to the next part. This questionnaire is anonymous, and you can choose to withdraw from participating at any time.

Participant number: ____

Gender: [Male] [Female] [Other]

When you were born in month/year: ___/______

Years of school/education: ____

Native speaker of Norwegian: [yes] [no]

Part two

Complete the sentences

In the following survey, you will be presented with incomplete sentences. Please fill in the blanks to form complete sentences.

Do not overthink your answer. Just write the first thing that comes to your mind. Also, please avoid long phrases: keep your answer short.

For every sentence, please state whether you are sure or unsure that the sentence is grammatically correct, or if you do not know the word/sentence.

Example answer:

Robert fixed...

his bicycle.

She annoyed...

[ ]

The complete list of sentences is as follows:

		Your answer	*Sure*	*Unsure*	*Do not know the word*
She	admired				
She	annoyed				
The board	approved				
The policeman	arrested				
The dog	ate				
Mila	boiled				
Robert	borrowed				
She	bought				
He	broke				
She	brushed				
Nancy	buckled				
Sarah	buried				
They	carried				
He	changed				
He	collected				
Kevin	cut				
Harry	decorated				
She	delivered				
She	dropped				
He	emptied				
Lily	entertained				
I	examined				
Emma	extinguished				
David	filled				
James	fired				
The woman	fried				
She	frightened				
Linda	ground				
Olivia	guarded				
Karen	hammered				

		Your answer	*Sure*	*Unsure*	*Do not know the word*
He	hunted				
Jennifer	impressed				
She	ironed				
He	itched				
Chris	judged				
She	juggled				
She	kicked				
They	knitted				
She	knocked				
Paul	landed				
He	licked				
Andrew	lit				
Mary	loaded				
She	married				
Mark	measured				
He	memorized				
George	mended				
Laura	milked				
Edward	missed				
He	moved				
He	obeyed				
Thomas	ordered				
He	owned				
Bob	painted				
Roger	pinched				
Ruth	planted				
She	played				
She	pushed				
She	read				
The man	reduced				
Charles	removed				
He	rubbed				
Ronald	scratched				
John	served				

		Your answer	*Sure*	*Unsure*	*Do not know the word*
He	sipped				
He	squeezed				
They	stroked				
She	threw				
Hannah	tied				
He	tightened				
Brenda	twisted				
Susan	typed				
Sandra	unfastened				
She	unlocked				
Lisa	visited				
He	washed				
The crowd	watched				
He	watered				
He	wore				
They	worshipped				

Appendix E – WMTB-sheet

TESTNUMMER:___ DATO:________

1. Digit Recall

	Practice List	Response	Score (1 or 0)
P1	2		
P2	1 5		
P3	7 4 8		

Span	List	Response	Score (1 or 0)
1	4		
	9		
	7		
	3		
	8		
	1		
2	6 2		
	4 9		
	9 1		
	3 8		
	7 4		
	2 5		
3	4 8 3		
	2 6 1		
	7 4 3		
	3 7 6		
	1 8 4		
	6 9 4		
4	5 9 2 6		
	3 1 7 4		
	2 8 5 1		
	9 6 2 7		
	8 5 1 4		
	7 2 8 6		

Span	List	Response	Score (1 or 0)
5	6 1 4 2 5		
	3 2 8 4 1		
	7 4 2 5 9		
	3 7 6 1 4		
	4 9 2 5 7		
	1 5 8 3 6		
6	8 3 9 2 5 1		
	2 9 5 7 1 3		
	1 6 2 5 9 4		
	7 5 9 2 6 4		
	6 8 1 3 7 5		
	5 2 7 3 8 6		
7	4 9 6 3 1 5 2		
	5 8 1 9 2 6 4		
	9 3 7 5 2 8 1		
	2 5 6 3 8 1 4		
	8 3 5 9 1 7 2		
	6 2 8 7 4 9 3		
8	1 7 5 6 9 3 2 4		
	9 2 8 1 7 4 6 3		
	8 6 4 2 9 3 1 7		
	6 1 3 7 2 9 4 8		
	5 2 9 4 7 3 1 6		
	3 8 4 1 5 2 7 9		
9	7 1 6 3 8 4 2 9 5		
	4 6 9 2 8 1 5 7 3		
	9 5 7 3 6 8 4 2 1		
	2 9 4 7 1 6 8 3 7		
	8 3 9 4 7 1 5 2 6		
	1 5 3 9 4 7 2 8 6		
		Trials Correct	
		Span	

9. Backward Digit Recall

	Practice List	Response	Score (1 or 0)
P1	2 3		
P2	5 4		
P3	3 4 5		
P4	5 2 4		

Span	List	Response	Score (1 or 0)
2	2 7		
	5 9		
	3 1		
	9 7		
	4 6		
	8 4		
3	8 1 4		
	6 3 7		
	4 6 2		
	2 5 9		
	7 3 5		
	9 4 3		
4	2 7 1 4		
	5 2 7 3		
	6 3 8 4		
	1 5 4 9		
	9 6 5 8		
	8 1 6 2		

Span	List	Response	Score (1 or 0)
5	2 1 4 9 8		
	5 7 1 4 2		
	2 7 4 6 3		
	9 5 1 4 2		
	3 5 8 2 6		
	4 6 3 1 5		
6	5 2 1 7 9 3		
	2 7 6 3 8 5		
	4 8 3 5 2 7		
	8 5 2 9 1 3		
	1 9 5 8 2 4		
	6 1 3 9 5 2		
7	8 3 5 2 9 4 1		
	6 3 1 9 4 7 5		
	5 8 7 2 4 9 3		
	7 9 2 6 1 9 3		
	8 5 2 4 9 3 6		
	9 6 2 8 1 4 7		
		Trials Correct	
		Span	

Appendix F – Consent form

Forespørsel om deltagelse i forskningsprosjekt
"Andrespråksbrukeres behandling av engelske verb"

Bakgrunn og formål
Jeg er en masterstudent ved Institutt for språk og litteratur ved NTNU som jobber med et forskningsprosjekt, der jeg trenger deltagere fra både videregående skole/ungdomsskole og universitetet. Prosjektet har fokus på tilegnelse av andrespråk, og målet med studiet er å undersøke hvordan norske fremmedspråksbrukere av engelsk prosesserer (hvordan hjernen behandler) engelske verb. Det er frivillig å delta, og jeg er svært takknemlig for alle som har muligheten til å være med.

Hva innebærer en deltagelse i studien?
Deltagerne vil være med på til sammen tre tester, som fordeles på to forskjellige tidspunkt i løpet av høsthalvåret 2016. I første runde vil deltagerne gjennomføre en kartleggingstest (kartlegging av arbeidsminnet) der de skal gjengi en rekke med tall/ord som blir lest opp. Test nummer to innebærer å fylle inn ord i ufullstendige setninger. Setningene vil vi presentert på et ark. Denne testen skal også gjennomføres i andre runde, som vil bli gjort nærmere jul.

Hva skjer med informasjonen om deg?
Alle personopplysninger (alder i måned/år, kjønn) vil bli behandlet konfidensielt. Alle deltagere får et deltagernummer som vil knytte navnene til resultatene. Det vil kun være lærer/veileder som har tilgang til listen som knytter navnene til deltagernummeret, og denne skal lagres utilgjengelig for uvedkommende. Enkeltpersoner vil ikke kunne gjenkjennes i publikasjonen. Prosjektet skal etter planen avsluttes i juni 2017. Personopplysninger vil da slettes, slik at datamaterialet er anonymisert.

Frivillig deltagelse
Det er frivillig å delta i studien, og du kan når som helst trekke ditt samtykke uten å oppgi noen grunn. Dersom du trekker deg, vil alle opplysninger om deg bli anonymisert. Dersom du har spørsmål til studien, ta kontakt med:

Mona Langeng Hammerås, masterstudent ved Institutt for språk og litteratur, NTNU.
E-post: monalh@stud.ntnu.no
Mila Vulchanova, professor ved Institutt for språk og litteratur, NTNU.
E-post: mila.vulchanova@ntnu.no
Giosué Baggio, førsteamanuensis ved Institutt for språk og litteratur, NTNU.
E-post: giosue.baggio@ntnu.no

Studien er meldt til Personvernombudet for forskning, Norsk samfunnsvitenskapelig datatjeneste AS.

Samtykke fra forelder/foresatte til deltagelse i studien
Jeg gir tillatelse til at min sønn/datter, __________________________, kan delta i studien.

(Signatur fra forelder/foresatte til prosjektdeltager, dato)