Roger Johnsen

Sensitivity to second language argument structure

An experimental study with Norwegian learners of English

Master's thesis in English – Teacher Training Program

Trondheim, May 2016

Supervisors: Mila Dimitrova-Vulchanova & Giosuè Baggio

Norwegian University of Science and Technology
Faculty of Humanities
Department of Language and Literature

Abstract

The aim of this master's thesis is to show what level of sensitivity Norwegian speakers of English display to the argument structure of English verbs. Quantitative experimental research was carried out in a group of 25 Norwegian upper secondary school pupils at the age of 16. An eye tracking experiment was conducted with English stimuli sentences and images on a screen either corresponding to the object of the sentence, a phonological competitor, or distractors. By processing native speaker data, the verbs of the stimuli sentences were grouped by different levels of constraint. The results showed that verbs with the highest level of constraint attracted higher gaze proportions for their corresponding object image than verbs of lower levels of constraint. This pattern was contingent on high scoring results in a grammar test occurring before the eye tracking experiment, indicating that the sensitivity develops at a rather high level of second language grammar aptitude. In an experimental condition where the target images were absent on the screen, the phonological competitor attracted higher gaze proportions, but the constraint of the verb did not influence gaze proportions.

Preface

Back in the autumn of 2014 I was faced with choosing what field to specialise in – linguistics, literature or civilization studies. Out of all the lower level English subjects I had studied so far, I found linguistics to be the hardest and most challenging. I felt that a firmer understanding of grammar and language acquisition would be a welcome preparation for the task of teaching English, which is why I ended choosing linguistics for my specialisation. The subjects I have studied since and the work I have put into this thesis has given me a firmer understanding of what language is, and I believe I will be a better teacher for it. Working with a long term project like this has taught me a few things about planning and organising that I believe will be useful experiences to carry with me to the classroom. At times it has been fun and rewarding; at times it has been frustrating.

I want to express my gratitude to the people whose helpful input has shaped the final outcome of the project. First of all, thank you to my supervisors, Professor Mila Dimitrova-Vulchanova and Associate Professor Giosuè Baggio, for their valuable help in shaping the project and providing feedback along the way. Thank you to my research partner Anders Schärer Reine for the fruitful cooperation we have shared these 12 months. Thank you to Martina Agnese Pentrella for helping us record our stimuli with her native English voice. Not least, I am very grateful to Hendrik Bertil Eshuis for all his help with coding the experimental design and helping us make sense of the data that came out the other end. Without his help, there would have been no thesis.

Finally, I am grateful to the teachers and professors who helped us organise and gather participants for both the native speaker experiment and the second language experiment, and to all the people who agreed to participate.

Roger Johnsen,
Trondheim, May 2016

Table of contents

List of figures

List of tables

1. Introduction

Sensitivity to the argument structure of verbs plays a predictive part in language processing. This has been demonstrated by previous research. Altmann and Kamide (1999) were the first to find that verbs would more quickly elicit predictive looks towards the appropriate object of that verb when only one item on a screen was a feasible object of the verb compared to verbs that could possibly select several of the items as its object. Kamide, Altmann, and Haywood (2003) found similar predictive behaviour based on argument structure in several other types of constructions. To the best of my knowledge, previous work within this area has been exclusively carried out in native speaker populations, and as such there is little research into the sensitivity to second language argument structure in second language users. The aim of the present thesis is to shed light on this issue by carrying out an eye-tracking experiment in a Norwegian participant group whose second language is English.

The research question that underlay the project was whether and to what extent native Norwegian speakers whose second language is English are sensitive to the argument structure of English verbs, and will exhibit this sensitivity through predictive behaviour in a language processing experiment.

The experimental design was partially modelled after the design of an experiment carried out by Brock, Norbury, Einav, and Nation (2008) in a group of adolescents with autism spectrum disorder. Their study had participants sit in front of a computer screen with a built-in eye tracker, and listen to recorded sentences while observing whether any one of four images on the screen matched any words in the auditory input. One image represented the object; one image represented a competitor with an identical onset syllable; two images were unrelated distractors. The stimuli sentences would have either a constrained verb (e.g. *drink*) or a neutral verb (i.e. *choose*), and the researchers measured gaze proportions towards an item corresponding to the object of the sentence from the onset of the verb. They found that both autistic participants and language-matched peers looked more towards the appropriate object than other objects well before the onset of the target word in the auditory input when hearing constrained verbs. They also found that participants looked more at a competitor object sharing the onset syllable than at phonologically unrelated objects when the verb was neutral.

The present study utilised a similar experimental design, but used a different approach to the stimuli. Three levels of verb constraint were created by looking at native speaker data for each verb, and a forth baseline category containing presumably neutral verbs (i.e. *choose*, *select*, *pick*, *take*) was also included. The aim was to see whether the level of constraint would

affect the gaze behaviour by way of increased looks towards the appropriate object for verbs that were more constrained in their argument structure.

The research presented in this thesis was the joint efforts of two master's students. We designed and carried out the experiment together, and tested two groups of different age and English proficiency: one group of 16 and 17-year-old pupils in their first year at a public upper secondary school (videregående skole) and one group of young adult university students having studied English for more than one year at university level. Each of us then processed the data for one of these groups in our respective theses; the present thesis looks at the data of the younger group whose English proficiency level is lower. As such, this thesis only offers half the picture, while a comparison across the two groups will provide a fuller understanding of the development of sensitivity to argument structure in second language users. Such a comparison is a planned extension of the project, but is not treated in the present thesis, and the conclusions offered here must be regarded as preliminary.

Our expectations were that both groups would display a certain sensitivity to argument structure, by showing increased gazes towards the appropriate objects when the level of verb constraint was higher and the target was present on the screen. In a target-absent condition, where the appropriate object was not present on the screen, but an object representing a cohort competitor with an identical onset syllable was present along with three unrelated distractors, our expectation was to find a phonological effect in the form of increased gazes towards the competitor. We also expected this effect to diminish with the increasing constraint of the verb, such that constrained verbs deterred gazes towards a phonologically related but semantically inappropriate object.

2. Theory

2.1. Verb meaning and argument structure

The theoretical background for this thesis is found within the field of lexical semantics, which is "the study of how and what the words of a language denote" (Pustejovsky, 1995, p. 1). The main assumption that underlies the thesis, is that the lexical meaning of a verb predicts the possible structures that this verb appears in at the level of syntax. This assumption guided Levin when she made preliminary attempts to systematise the facets of verb behaviour. "[T]he behavior of a verb, particularly with respect to the expression and interpretation of its arguments, is to a large extent determined by its meaning" (Levin, 1993, p. 1). In order to examine this idea, we need an understanding of such terms as *lexicon*, *argument, thematic role, theta grid, argument structure, event type* and *context*.

The *lexicon* is the "vocabulary" of a language, and the *mental lexicon* refers to this system as it is represented in the mind. Many views on how the lexicon is organised have been formulated, the most famous of which according to Levin (1993) stems from Bloomfield, who claimed that the lexicon is a list of irregularities and idiosyncrasies associated with each word; an appendix of the grammar. This idea aligns with a popular view within generative grammar that the lexicon should only consist of the minimum necessary information about each lexical item (p. 1). Levin challenges this idea, and claims that the knowledge demonstrated by speakers "suggest that there is more to lexical knowledge than knowledge of word-specific properties" (Levin, 1993, p. 1). Levin's examples concern the properties of verbs.

A sentence describes a situation, the number of entities involved and the roles of these entities, and all this information is assumed to be encoded in the lexical entry for the verb (Koenig, Mauner, & Bienvenue, 2003, pp. 67-68). Consider the sentence in (1):

(1) Tommy cleaned the window.

The sentence identifies two entities, *Tommy* and *the window*, related by the action described by the verb *clean*. The entities are portrayed in specific roles: *Tommy* carries out the action and *the window* is acted upon. These entities whose presence and participation in the situation are associated with particular verbs are typically called *arguments* (Koenig et al., 2003, p. 68). It is common to say that a verb selects its arguments, and assigns different roles to them. These roles have been labelled in different ways, but I will keep to Dowty's tradition of calling them *thematic roles* (Dowty, 1991, p. 547), *theta roles* for short. Dowty claims that a

lack of consensus about what thematic roles actually are is notable in the linguistic community, and that no complete list of roles has ever been proposed, and while they belong to the syntax-semantics interface, appeals to them can be a confusion of notions from syntactic, semantic and pragmatic domains (Dowty, 1991, pp. 547-548). Even so, there are some commonly used thematic roles that will suffice for the purpose of the present thesis. Going back to the sentence in (1), *Tommy* is the initiator of the action, and he is acting with volition. He is the given the role [AGENT]. *The window* is the entity undergoing the effects of the action, and is given the role of [PATIENT]. Some other common thematic roles are [THEME] (an entity moved by an action, or whose location is described), [EXPERIENCER] (an entity that is aware of the action or state described, but not in control of the action or state) and [BENEFICIARY] (an entity for whose benefit the action was performed) (Saeed, 2009, pp. 153-154). Several other thematic roles have been proposed by many writers, but I will not pursue this issue further for the present purposes.

Verbs have certain requirements for their thematic roles. Not only do they specify how many arguments they require (i.e. whether the verb is intransitive, transitive or ditransitive) but also what thematic roles its arguments hold. According to Saeed (2009) this listing of thematic roles is usually called *thematic role grid*, or *theta-grid*, and can be listed like the the example in (2).

(2) **put** V: <AGENT, THEME, LOCATION>

The entry in (2) tells us that *put* is a three-argument verb (ditransitive), and that the thematic roles of the arguments can be [AGENT], [THEME] and [LOCATION]. The theta-grid predicts that the verb *put* can be used to form a sentence like *John put the book on the shelf*, with *John* becoming the [AGENT], *the book* becoming the [THEME] and *the shelf* becoming the [LOCATION] (Saeed, 2009, p. 160). Going back to Levin's claim that verbs show particularly complex properties that need to be accounted for in the lexicon, the theta-grid illustrates part of what she is referring to. Since the information encoded in the theta-grid is part of a speaker's semantic knowledge about the verb, we should expect it to be part of the lexical information stored in the lexicon.

What has been presented so far about verbs and their arguments can be attributed to the term *argument structure.* We say that a verb has argument structure, and this feature encompasses both the semantic properties expressed in the theta-grid – the number of arguments that the verb selects and their semantic relation to the verb – and the syntactic properties related to where and how the different thematic roles are realised in the syntax. As seen, the selectional requirements of the verb are closely related to the verb meaning. The

syntactic realisation of a theta role is not necessarily locked. An [AGENT] usually appear as the subject of the sentence, but certain thematic roles can be realised in several different syntactic positions. An example is *locative alternation*, which captures the fact that certain verbs can express their arguments in two different ways. Consider examples (3a) and (3b) from Levin (1993).

(3a) Sharon sprayed water on the plants.

(3b) Sharon sprayed the plants with water.

These sentences are both acceptable, but verbs that are closely related to "spray" do not always allow the same alternation. Consider "cover", which is similar to "spray" in meaning, in that both verbs relate to covering surfaces. Both verbs can select a location argument as direct object (the *with* variant), in which case the location is understood to be completely affected by the action (Levin, 1993, pp. 118-120). "Cover" does not allow the same alternation, as seen in (4a) and (4b).

(4a) *Monica covered a blanket over the baby.

(4b) Monica covered the baby with a blanket.

Levin argues that native speakers are able to make extremely subtle judgements concerning what syntactic expressions are allowed by each verb, this locative alternation being one such example. Another such syntactic judgement is whether a verb may participate in *transitivity alternations*: alternating between intransitive and transitive use (Levin, 1993, pp. 2-3). She argues that the key to the ability of making such judgements about the behaviour of a verb is in its meaning. Levin points to an example by Hale and Keyser, who considered the verb *gally*, which is a little known English whaling term used in sentences like *The sailors gallied the whales*. Being unfamiliar with the verb, we could presume that it means something like "see" or something like "frighten". A native speaker with the first meaning in mind will make the following judgements about the verb's syntactic behaviour: it does not allow the *middle voice* construction *Whales gally easily* (cf. *Whales see easily). Meanwhile, the second meaning (and incidentally the correct one) would render the construction perfectly acceptable (cf. Whales frighten easily). Hence, particular syntactic properties are associated with verbs of certain semantic types (Levin, 1993, pp. 4-5). The conclusion to be drawn from Levin's work is that lexical semantic features play a determining role in the syntax of the argument structure (Stringer, 2010, p. 109).

Dimitrova-Vulchanova and Dekova (2007) argue that the syntactic behaviour of a verb is also closely related to the type of event that the verb lexicalises. Within the *Sign Model*, the situation lexicalised by the verb can be captured by several different dimensions where

elements are assigned values on these dimensions. One such dimension is *Force*, capturing situations where emission of physical force is happening, and values assigned to this dimension can be *Source* (the entity from whom the force is released), *Source extension* (the specific part of the source entity performing the action) and *Absorber* or *Limit* (the entity affected by the force). To illustrate how event types influence morphosyntactic behaviour we can look at *Contact* situations, which are exclusively characterised by the contact obtained by participants, such that they cannot be omitted from realisation in the syntax. Ungrammaticality ensues when the direct object of *slapped* is not overtly realised (cf. *Angel slapped). The *Absorber* in the *Force* dimension can also be coindexed with *Monodeveloper*, representing a different dimension – *Monodevelopment*, in situations where the absorber undergoes a unidirectional process that for instance changes its integrity as a result of the *Force*, such as a screen ending up scratched as a result of someone tapping it. The English verb *scratch* can also lexicalise a different event type where there is no Monodevelopment, such as someone scratching their head. Some languages, like Bulgarian, would lexicalise these different situations using different verbs (Dimitrova-Vulchanova & Dekova, 2007, pp. 37-44). When a single verb can lexicalise several different event types, context helps disambiguate the verb meaning.

The *context* in which an utterance is made contributes to the linguistic encoding of the situation. However, there are wide-ranging definitions of what context comprises, from a narrow focus which counts only the linguistic information co-occurring with the utterance, to a wide focus which could include the discourse, or even facts or circumstances in the knowledge periphery of an event (Dimitrova-Vulchanova & Weisgerber, 2007, p. 52). The importance of context to determine the meaning of the situation described by the verb can be illustrated by the examples in (5), (6) and (7) from Dimitrova-Vulchanova and Weisgerber.

(5) Peter climbed the ladder.

(6) The airplane climbed to 10,000 feet.

(7) The balloon climbed.

An interpretation of the motion situations described by the verb *climb* in these sentences relies on knowledge about the moving object, the ground object and their interaction on a path. Dimitrova-Vulchanova and Weisgerber (2007) argue for a wide definition of 'context' which includes "relevant knowledge available in the same place and time as an event" (p. 55). They argue that relevant knowledge is likely activated, not by all knowledge about a situation trying to enter into the processing, but rather by a lexical entry containing open slots of underdetermined information, such that the verb "knows what it does not know". The verb

then searches for knowledge that has been implicit, such as references to implicit objects or paths (Dimitrova-Vulchanova & Weisgerber, 2007, p. 70).

The theoretical background presented so far argues for an understanding of morpho-syntactic behaviour of verbs as something encoded in their lexical meaning. The following will show how language users use this information in a predictive manner in processing, and how this can be studied in the field of psycholinguistics.

2.2. The visual world paradigm

The *visual world paradigm* refers to a research paradigm and a methodology pioneered by Cooper in 1974, then further developed by Tanenhaus and colleagues from 1995 (Huettig, Rommers, & Meyer, 2010, p. 152). The common methodology of the paradigm involves presenting participants with oral linguistic input while they see a visual scene comprised of either physical objects or images on a screen that are relevant to the spoken input. Researchers then monitor the eye movements during linguistic input using a camera, and analyse these movements to study how linguistic input affects the gaze pattern of the participant. Cooper (1974) found that when participants were instructed to look anywhere, they tended to gaze towards objects that were mentioned or associated with the linguistic input that was presented to them orally, and that these eye-movements were closely time-locked to the linguistic input, with 90% of gaze fixations on the critical objects happening when the word was spoken or within 200 ms after word offset. It has since been established that there is a latency between when saccadic eye movement is programmed and when a fixation occurs, and Allopenna, Magnuson, and Tanenhaus (1998) found that fixation on relevant objects start about 200 ms after the onset of its associated target word (p. 428).

Because of the time-locked relationship between spoken word input and eye movements, the recorded gaze data provide a sensitive and non-disruptive measure of language comprehension. Because modern equipment has a high gaze data recording frequency, the method provides very good temporal resolution. Lexical processing can be studied in the context of ongoing speech, without the need to interrupt this speech to ask participants questions about comprehension or metalinguistic judgements (Allopenna et al., 1998, p. 421).

Some issues can be raised against findings within the paradigm. As Huettig et al. (2010) point out in their critical evaluation, the link between eye movement and language processing is indirect. The gaze of a participant indicates the focus of their visual attention, which is not only dependent on spoken input, but also on other elements, such as the

participant's working memory representation of the objects, higher-level inference processes, and their understanding of and compliance with any task they are given. Processing is based both on the linguistic and the visual input, and the visual context may limit the mental processes that would otherwise occur from spoken word input alone. Linguistic input is compared directly to the limited set of possibilities set up by the visual context, rather than the natural case where the linguistic input must both generate possibilities and facilitate selection between these possibilities. There is also a possibility that results are biased towards greater processing speeds, as little linguistic information will be needed to distinguish between a very limited number of alternatives presented in the visual context. Regardless of these and other objections, Huettig et al. conclude that the visual world paradigm is very well suited for studying and assessing how the mind processes utterances about objects and events that we see, and particularly the interplay of language, vision, memory and attention (Huettig et al., 2010, pp. 166-167).

The notion of *competitor* is relevant in speech processing, and will be significant in the present thesis. Acoustic input is processed without delay on a word-by-word basis while it is being received. When a person hears an acoustic string (i.e. the first the part of a word, then the whole word, then a string of words) all the lexical items that are momentarily compatible with the input will be activated initially, all competing for recognition. As more input is received, certain alternatives become incompatible, and are deactivated. As long as these alternatives are not yet excluded, they act as phonological competitors (Altmann, 1997, p. 9). Compelling evidence for this hypothesis is found in many studies; for instance, it has been found that words with several syllables are recognised at the point where the phonetic input is no longer consistent with other lexical candidates, and that the reaction time for spoken words depends on the number and frequency of other words that differ by only one phoneme. In the *cohort model* developed by Marslen-Wilson and colleages, all the lexical candidates that are activated by the onset of a word – all words that share onset syllables (like "beetle" and "beaker"), comprise a cohort that competes for recognition as long as they are compatible with the input (Allopenna et al., 1998, pp. 419-420). Allopenna et al. (1998) found the effect of the cohort competitor (i.e. a higher probability of fixating on the competitor than unrelated objects) to start around 200 ms post target onset. This was the same time as the effect of the target referent became apparent. The effect of the competitor reached its peak around 500 ms post target onset, before diminishing (pp. 428-429).

2.3. Acquiring a sensitivity to argument structure

One of the major insights established within the visual world paradigm is that language processing is driven by a predictive relationship between a verb, its arguments and the real-world context in which they occur. As such, language processing must be sensitive to the argument structure of verbs presented in section 2.1. The first experimental research that found evidence for this hypothesis was conducted by Altmann and Kamide (1999). They found that when participants were presented with the sentence *The boy will eat the cake,* which contains the semantically constraining verb "eat", they would more quickly activate the appropriate filler for the [THEME] argument slot, "cake", than when presented with the sentence *The boy will move the cake,* where the verb is open to selecting other, non-edible fillers of [THEME] argument slots. Evidence for this activation was a higher probability of looking towards an image of a cake when hearing the former sentence than when hearing the latter. Kamide et al. (2003) expanded on the insight from Altmann and Kamide's experiment with monotransitive verbs, and found that the most appropriate post-verbal [GOAL] argument slot filler in ditransitive sentences was anticipated at the point of the verb. In the sentence *The woman will spread the butter on the bread,* the verb "spread" led participants to anticipate the [GOAL] "the bread" rather than "the man". Opposite, in the sentence *The woman will slide the butter to the man,* the verb "slide" led to the anticipation of "the man" as [GOAL]. Further, the same researchers found that a pre-verbal argument ([AGENT]) in combination with the information conveyed by the verb could constrain anticipation for subsecquent [THEME] argument slot filler. The sentence fragment *The man will ride…* increased fixations to a motorbike, while *The girl will ride…* increased fixations to a carousel. Finally, in an experiment with the Japanese language, where verbs appear at the end of the construction (head-final), the combination of the [AGENT] argument slot filler "waitress" and [GOAL] argument slot filler "customer" led to the anticipation of the [THEME] argument slot filler "hamburger" before the verb "bring" was uttered.

These experimental results illustrate the sensitivity to argument structure that first language (L1) users display. Research in language acquisition has found that sensitivity to the behaviour of verbs is acquired at an early age. Brooks and Tomasello (1999) investigated how children acquire constraints on their syntactic constructions, and asserted that the development of sentence-level syntactic constructions follows a U-shaped curve. Very young children rarely use a newly acquired verb in a structure that differs from the ones they have already heard it used in (such as taking a verb that they have only heard in transitive

constructions and using it in an intransitive one, or vice versa), suggesting that the earliest constructions are specific to lexical items rather than abstract word classes. They then gradually generalize across these lexical items, until there are instances of overgeneralisations (e.g. *Don't giggle me!*). Eventually they acquire adult-like constructions. The results of the study indicated that by about 2.5 years of age children had acquired abstract transitive and intransitive constructions that supported productive usages with newly acquired verbs, and at 4.0 to 4.5 years of age they started constraining their use of constructions to adult-like ranges on the basis of semantic classes of verbs (Brooks & Tomasello, 1999, p. 736).

In the domain of second language acquisition, Stringer (2010) argues that *lexical relativity* accounts for the difficulty facing L2 learners when acquiring the argument structure of verbs in their L2. The principle of lexical relativity states that any lexical concept in a language is determined relative to other lexical items, and accounts for the fact that it is extremely difficult to find true cross-linguistic matches when comparing items of two languages. The meaning of a word is determined not only by its relationship to the concept it denotes, but also to other words in the same language. One example that serves to illustrate this point is how the French word *mouton* corresponds both to English *sheep* and *mutton*, which respectively denote the animal and its meat. Because English has separate terms for the two, the concept denoted by *sheep* does not fully match the concept denoted by the French *mouton*.

Lexical relativity extends to verbs and their argument structure, and is according to Stringer fundamental to our understanding of the initial state of L2 learners and their developmental path (Stringer, 2010, p. 105). To illustrate this idea with an example, the Korean verbs that match English ground-oriented locative verbs most closely often allow locative alternation, unlike the English ones (fill the glass with water/*fill the water into the glass). Korean learners of English tend to wrongly assume lexical equivalence, and consequently produce errors in English (Schwartz, Dekydtspotter, & Sprouse, 2003). Stringer makes the case, based on the Full Transfer/Full Access model developed by Schwartz and Sprouse, that the L1 lexicon is the initial state for L2 lexical acquisition. In other words, the L2 learner initially transfers all the features of an item, both semantic and syntactic, from their L1 to the perceived L2 analogues – the lexical items that most closely matches their native language items. However, experimental findings from Dekydtspotter, Schwartz, Sprouse, and Bullock (2008) indicate that L2 learners eventually acquire native-like judgements. If the findings are generalizable, this development parallels that of first language acquisition;

children also overgeneralize alternation patterns and later develop a more restrictive grammar, despite no negative evidence in the input (Pinker, 1989).

Treffers-Daller and Calude (2015) showed that word frequency is essential for acquisition of L2 structures, and that L2 learners are sensitive to word frequency in the language they are learning. In their study, frequent usage was shown to be a key determinant of acquisition of French motion verbs for English learners. English learners of French have difficulty learning motion verbs, because in French the path by which an object is moving is often encoded in the verb, as in *traverser* ("to cross"), while in English the path is typically expressed in a particle associated to the verb, like *across* (Treffers-Daller & Calude, 2015, p. 8). Treffers-Daller and Calude found that statistical learning caused L2 motion event patterns to become entrenched and replace patterns from the L1. They also found that higher level learners were better able than intermediate level learners to match native-speaker frequency of usage of motion verbs.

The findings from Dekydtspotter et al. showing that Korean L2 users eventually acquire native-like judgements with regards to English locative verbs and the ones from Treffers-Daller and Calude showing sensitivity to word frequency in the L2 both indicate that proficient L2 users eventually develop an advanced understanding of argument structure in their L2. To what extent they are sensitive to argument structure in their language processing, like Altmann and colleagues found L1 users to be, is unknown. At what stage in the L2 development such a sensitivity should emerge is also unknown. The aim of this thesis is to shed light on these questions by submitting native Norwegian learners of English to experimental testing within the visual world paradigm.

3. Methods

The research approach of the present study is quantitative, which allows for a comparison of results across a rather large number of participants. The experimental design consisted of grammar and vocabulary tests, and an eye tracking procedure where participants listened to English sentences and watched images on a computer screen, while gaze patterns were recorded by the eye tracker. The gaze behaviour of the participant group was measured across several variables (four different verb categories, as well as a target-present and a target-absent condition), and the resulting numerical data was statistically analysed. In this section I will present a discussion of the preparatory stage of the project, the participants of the experiment, the materials and procedure used in the experiment.

3.1. Preparatory stage

An important aspect of the present study was to examine how the level of constraint enforced by the argument structure of a verb on its object would affect the gaze data. The expectation was that a more constraining verb would aid participants and increase the probability of looks towards the target object compared to a relatively less constraining verb. The idea was to include three categories of verbs with three different levels of constraint. As a basis for this categorisation, 77 common English verbs that intermediate L2 users were expected to know were used as a starting point. Verbs initially thought to be highly constraining as well as verbs considered very open with regards to argument selection were included in this selection. For instance, a verb like *milk* was assumed to select an NP for its [PATIENT] argument slot that is a mammal in its most salient use, commonly a cow. The verb *buckle* was assumed to often select for its [THEME] argument slot a physical object that has a buckling mechanism (e.g. belts or shoes). Contrast these to the verb *remove*, which was expected to be very open to selecting almost any entity as a [THEME] argument.

In order to separate these verbs into three categories of ever increasing constraint, an online survey to elicit natural native speaker uses of these verbs was prepared. In the survey, participants were presented with 77 incomplete sentences containing a subject (a common English name or a pronoun) and a verb. Participants were instructed to complete the sentence by writing the first thing that came to mind, and to keep their answers short. This effectively provided a range of NP fillers of the [PATIENT]/[THEME] argument slots for each verb. By evoking the "first whims" of participants, the expectation was that the resulting data would reflect not only what possible direct objects help form well-formed utterances, but also the fact that certain fillers of argument slots are strongly associated with a given verb through

salient prevalence in language use. The survey and associated information sheet and consent form (see appendices C and D) was reviewed and approved by the Ethics committee at a British university and distributed to native English speaking students at the same university. Participation in the survey was voluntary but encouraged by a professor at the university. 105 students participated in the survey. Their ages ranged from 18 to 45 (mean age 19.75), and the gender distribution was 86% female and 14% male.

With 105 individual answers to each subject-verb string, the following procedure was used to distribute the verbs into categories of fairly equal size. For each verb, if three unique answers or fewer collectively constituted at least 55% of all the given answers, the verb was assigned to the most constrained category. If four to six unique answers constituted at least 55% of all the answers, the verb was assigned to the moderately constrained category. If seven or more unique answers were needed to constitute 55% of all answers, the verb fell into the least constrained category. When judging the uniqueness of an answer, only heads of noun phrases were counted, ignoring any prepositional phrases or adverbial phrases. Determiners ("a", "the", "his/her/their") were also ignored; as was number (singular or plural). As such, the answers "the cat", "his cats" and "a cat" would all count as the same answer – "cat".

The categorization resulted in 21 verbs in the least constrained category, 21 in the moderately constrained category, and 35 in the most constrained category. From each category, 16 verbs were selected for use in the experiment. A fourth category was constructed for use in the eye tracking experiment which contained the near-synonym verbs *choose*, *pick*, *select* and *take*, all of which were believed to be neutral with regards to argument selection. 24 instances of these sentences were included such that each object used in the trial was the target object of one sentence. This category functioned as a baseline against which the other categories could be compared. The auditory stimuli is further discussed in section 3.4.1.

3.2. Participants

The participant group comprised of 25 adolescents at the age of 16 years old. There were 15 female and 10 male participants. They were recruited from two VG1 classes in a public Norwegian upper secondary school. All participants were native Norwegian speakers with English as their L2. One participant reported that they had Spanish as an additional native language.

The school administration and the students' teachers agreed to having the experiment carried out during school hours, and all pupils in the two classes were encouraged to participate by the teachers, although it was emphasised that participation was voluntary. No

reward was offered. The study was reported to and approved by Norsk senter for forskningsdata (NSD). Given that the study did not collect sensitive personal information, and that participants were all 16 years old at the time of the experiment, they did not need parental consent to participate in this study. Participants signed a consent form approved by NSD (see appendix A). To prevent identification, each participant was assigned a participant number, and the link between participant names and numbers was only known to the teachers.

3.3. Procedure

The main part of the experiment consisted of an eye tracking procedure. Prior to this, participants filled out a questionnaire in paper form with questions concerning their linguistic background and other factors that were believed to potentially affect/explain their performance (see appendix B). Questions concerned reading and writing habits in their native language(s), self-assessment of English skills with regards to reading, writing, talking and listening, what other languages they knew, as well as other questions that were thought to be relevant, like habits for watching TV and playing computer games. Finally, several questions concerned whether they had any issues like hearing impairments, impaired eyesight or diagnoses that could impair language learning. Some questions were 'yes' or 'no' questions, others had participants choose between a range of five options from 'never' to 'every day'. Answers from this questionnaire led to exclusion of two participants: one participant reported English as an additional native language; one participant reported language impairments (such as dyslexia) as well as other diagnoses that could affect language acquisition (such as ADHD or autism).

After filling out the questionnaire, participants completed a grammar test and a vocabulary test on the internet using their personal computers. The grammar test (http://www.examenglish.com/leveltest/grammar_level_test.htm) consisted of 15 multiple choice tasks where the participant was presented with a sentence lacking one word, and a choice between four alternatives for what word to insert. Questions got easier or harder according to their responses. Their performance in the test led to an approximation of the participant's CEF proficiency level (A1 to C2). The vocabulary test (http://vocabulary.ugent.be/) was developed at the Center for Reading Research at Ghent University in Belgium, and consisted of a word recognition task. 100 letter sequences, some of which were existing English words (with American spelling) and some of which were non-words, were presented on the screen one by one, and participants had to press one of two buttons to indicate whether or not the letter sequence was an English word they knew.

Answering "yes" to a non-word affected the score negatively, and at the end they would get a result indicating approximately what percentage of the English words they knew. For each of the tests, the experimenters checked and wrote down the results to prevent participants from reporting false results.

The experimental design of the eye tracking procedure consisted of a setup where participants would sit in front of a computer screen with a built-in eye tracker (Tobii T120) that recorded gaze data at 120Hz sample rate. Participants wore headphones (Philips Cityscape Fixie) for auditory input and were requested to press one of two buttons on a game pad (Microsoft SideWinder Plug & Play Game Pad). Before the experiment started participants received oral instructions to look at the screen, listen to the sentences that were read aloud by the pre-recorded voice and see if any of the pictures on screen matched anything that was said in the sentence they heard. They were told to push one button for "yes" if there was a match, and another button for "no" in the opposite case. If uncertain, they could refrain from pushing any button. Before the first trial, the tracking of each participant's eye movements was calibrated. Between each trial they had to fixate for 150 ms on a point in the middle of the screen where a rotating figure was showing.

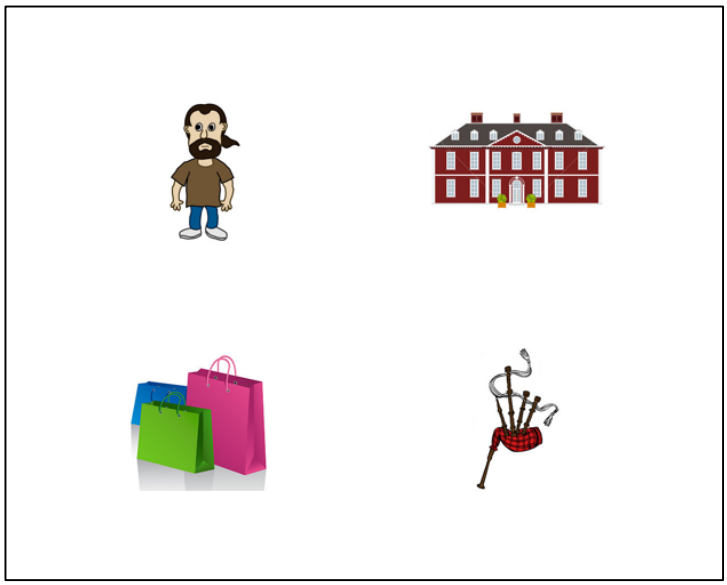


Figure 1: An example of a trial screen. The target image "man" is displayed on a white background along with the cohort competitor "mansion" and the two unrelated items "bags" and "bagpipe". Image positions were randomised.

At the start of each trial four images consisting of coloured illustrations were presented on a white background. Each image was confined within an *area of interest* (AOI) that took up 25% of the width and 25% of the height of the screen in a 2x2 setup with white space between them. Figure 1 shows an example of a trial screen. 1000 ms after the images appeared, the pre-recorded sentence started playing. This latency between visual and auditory stimuli is important, as it gives participants time to familiarise themselves with the objects and retrieve relevant visual representations about the objects, and increases likelihood of fixating on particular objects (Huettig et al., 2010, p. 153). If a participant pressed a button before the sentence had finished playing, their response was recorded, but the sentence played

to its end, and a short pause ensued with the images still displayed on the screen before the trial ended and the rotating fixation point appeared on the screen.

The auditory stimuli consisted of a sentence, digitally recorded at the Phonological lab at NTNU by a female native English speaker. Each sentence was five words long, and consisted of a subject (a common English name), verb, definite article, target noun, and adverb. The experiment consisted of 72 trials in a target-present condition and 72 in a target-absent condition. In the target-present condition the screen showed an image representing the target noun, a cohort competitor which referenced a noun that has an identical onset syllable to the target, and two unrelated distractors. In the target-absent condition the cohort competitor was shown along with three unrelated distractors. The screen positions of target, competitor and distractors were randomised for each trial. The trial order was semi-randomised (see section 3.4.2 for details). 24 images were used throughout the experiment, and distractor images acted as target and competitor in other trials, so that each image was shown an equal amount of times throughout the experiment in order to control for stimulus characteristics like word frequency or picture salience. The verb of each sentence was assigned to one of four different categories with regards to its selectional restriction ranging from least to most constraining. This verb categorisation is further explained in below.

3.4. Materials

3.4.1 Auditory stimuli

On the basis of the 48 verbs finally selected for inclusion in the three verb categories and the 24 baseline items, 72 stimuli sentences were constructed. As previously stated, each sentence was five words long, and consisted of a subject [AGENT], verb, definite article, object noun, and adverb. As subject of the sentences, the five names *Alex*, *Sam*, *Charlie*, *Jesse* and *Mary* were used. The first four can be argued to be ambiguous with regards to gender, while the latter is unambiguously female. For most sentences, the gender ambiguity of the name was chosen to avoid interference from any preconceived gender notions, although finding truly gender neutral names is challenging. The case can perhaps be made that Norwegian L2 users of English know most of the names primarily as male names from language input. As for sentences with the subject name Mary, these sentences would have an image representing “the man” either as its target, competitor or distractors, and given that participants were instructed to judge whether any word in the utterance matched any image on the screen, an image of a

man could be interpreted as the sentence subject, unless the subject was unambiguously female.

An adverb was included as the final word of each utterance to avoid any unwanted effects from having the target noun as the utterance-final word. From a selection of eight adverbs (carefully, happily, gently, regularly, accidentally, quickly, eagerly was hastily), the one considered most fitting to the context of the sentence was chosen.

As for the object noun of each sentence, it would be tempting to select the most common answer of each item in the L1 survey mentioned above. However, the experimental design relied on each target noun functioning as cohort competitor to a different object noun, and as distractor in other instances. In addition, each object functioned as the target for two different sentences, and by extension as competitor for two different ones. This allowed for the use of the same 24 images throughout all trials. As such, it was crucial that the onset syllable of each object noun was identical to the onset syllable of a different object noun. All these considerations constrained the possibility of what items to use as objects.

Another consideration made when constructing object nouns was avoiding that competitor and distractor nouns were feasible candidates for the object position of the sentences. Such a case would impair the effect of verb-stimulated looks towards the target noun, as the competitor and distractor nouns would be equally legitimate targets. This consideration applied to a larger extent to the verbs of the moderately and most constrained categories than the ones in the least constrained category, as it is the nature of verbs that would be characterised as "least constraining" with regards to argument selection that they can select a great variety of different nouns as legitimate objects. As such, it was difficult to avoid that the target nouns of some of the "least constraining" verbs appeared on the screen with cohort competitors and/or distractors that competed for selection even at the point of verb offset.

In order to comply with the preferences discussed so far, while still attempting to apply as much of the findings from the L1 survey as possible, object nouns were constructed that either matched one of the most frequent answers for each verb identically, or came as closely as possible with regards to semantic characteristics. For instance, the item *laces* was constructed for the verb *tie*, which had *shoelaces* as its most common answer. This object functioned as competitor to *laser-gun*, which was constructed from the frequent answer *gun* to the verb *fire*. Other instances complied even less with the L1 survey answers. For instance, *flour bags* was constructed as a feasible target for the verbs *deliver* and *empty*, in order to

function as a competitor to *flowers*, while L1 survey answers for neither *deliver* nor *empty* included *flour bags*. For a complete list of stimuli sentences, see appendix E.

3.4.2 Visual stimuli

As mentioned in section 3.3, the experiment trials featured four images consisting of coloured illustrations presented on a white background: a target item, a cohort competitor and two unrelated distractors. A crucial premise for studies involving language mediated eye movements is the association between the auditory input and the visual input on the screen. Participants need to perceive an immediate and implicit connection between for instance the target noun *lemon* and the image representing a lemon, otherwise the auditory input will not effectively mediate eye movements. Some previous studies within the visual world paradigm explicitly expressed this relationship between words and images by prefacing the experiment with a presentation and naming of all the items participants would see during the experiment (Allopenna et al., 1998, pp. 426-427). The present study had no such preface. In order to ensure that the image representing *back door*, *cattle*, *lemming* and *laces* were perceived as such and not as for instance *door*, *cows*, *hamster* and *shoelaces* respectively, which would be semantically similar, but render them inefficient as cohort competitors, all trials where these items functioned as target objects were presented to participants prior to trials where they acted as competitors. This was ensured by having the experiment run a group consisting of half the trials, including these, in a random order before running the the other half of the trials, which included sentences where the items where competitors in random order. By the time participants saw these images as competitors, they had already heard them referred to as *back door* rather than *door*, etc. For a complete overview of stimuli images, see appendix F.

3.5. Analysis

In order to carry out analysis on the data, gaze data as well as results from grammar and vocabulary tests were imported into the statistical analysis software *SPSS*. Answers from the language background questionnaire were ultimately not included in the analysis, as actual language test performances were assumed to be more relevant measures for comparison. For both the target-present and the target-absent condition, a critical time window was defined (see section 4.3 and 4.4), over which gaze proportions to the spaces of interest containing the target images were calculated. Trials with less than 25% registered gaze points (due to participants blinking or looking away from the screen, or the eye tracker not being able to register eye movement) were excluded from the analysis. 2.03% of the trials were excluded

this way. Analysis of variance (ANOVA) was applied to determine whether there was an effect of verb category on gaze proportions, and post-hoc pairwise comparisons were carried out to see whether gaze proportions for each verb category was significantly different from the others. Then an analysis of covariance (ANCOVA) was applied to determine if the language measures (the grammar and vocabulary results) could predict the effect that the verb category had on gaze proportions in the critical time window.

4. Results

4.1. Grammar and vocabulary results

The results of the grammar test ranged from A1 to C2 in accordance with the CEF proficiency levels. These scores were converted to numerical values (A1 = 1, A2 = 2, B1 = 3, B2 = 4, C1 = 5, C2 = 6) for the analysis. The mean grammar score of the group was 4.72 (std. deviation = 1.5948).

The results of the vocabulary test were given as a percentage of English words known. This value was converted to a numerical value. The mean vocabulary score was 33.44 (std. deviation = 13.1975). No significant correlation was found between participants' grammar and vocabulary scores (.281, p = .173).

4.2. Response accuracy and reaction time

Response accuracy reflects to what extent participants pressed the button indicating "yes" when the target object was present on screen, and "no" when it was absent. If a participant answered "yes" for all target-present trials and "no" for all target-absent trials, their accuracy proportion score would be 1.0. Reaction time (RT) indicates the time in milliseconds from sentence onset until the participant presses one of the buttons. Table 1 shows the mean accuracy and reaction times for items in each of the verb categories.

Table 1: Mean accuracy and reaction time across verb categories.

Target present?	Verb category	Mean accuracy	Std. deviation	Mean RT	Std. deviation
No	0 (baseline)	.8500	.06477	3194.10	317.14
	1 (least constraining)	.8850	.09319	3272.48	311.82
	2 (moderately constraining)	.8925	.09114	3348.64	297.01
	3 (most constraining)	.9000	.09021	3295.19	371.99
Yes	0 (baseline)	.8950	.07523	3085.03	295.44
	1 (least constraining)	.9200	.07108	3139.29	255.35
	2 (moderately constraining)	.9050	.08861	3306.92	272.51
	3 (most constraining)	.9200	.07971	3241.46	340.34

The results presented in Table 1 show that mean accuracy overall was fairly high, but that accuracy was slightly lower for the target-absent condition compared to the target-present

condition. Reaction time was slightly faster for the target-present condition compared to the target-absent condition, which indicates that participants were faster able to decide that an item referred to in the auditory input was present on the screen than to decide that none of the objects on the screen corresponded to anything in the auditory input.

Figure 2a shows mean accuracy across verb categories and target presence conditions.

Figure 2b shows mean reaction time.

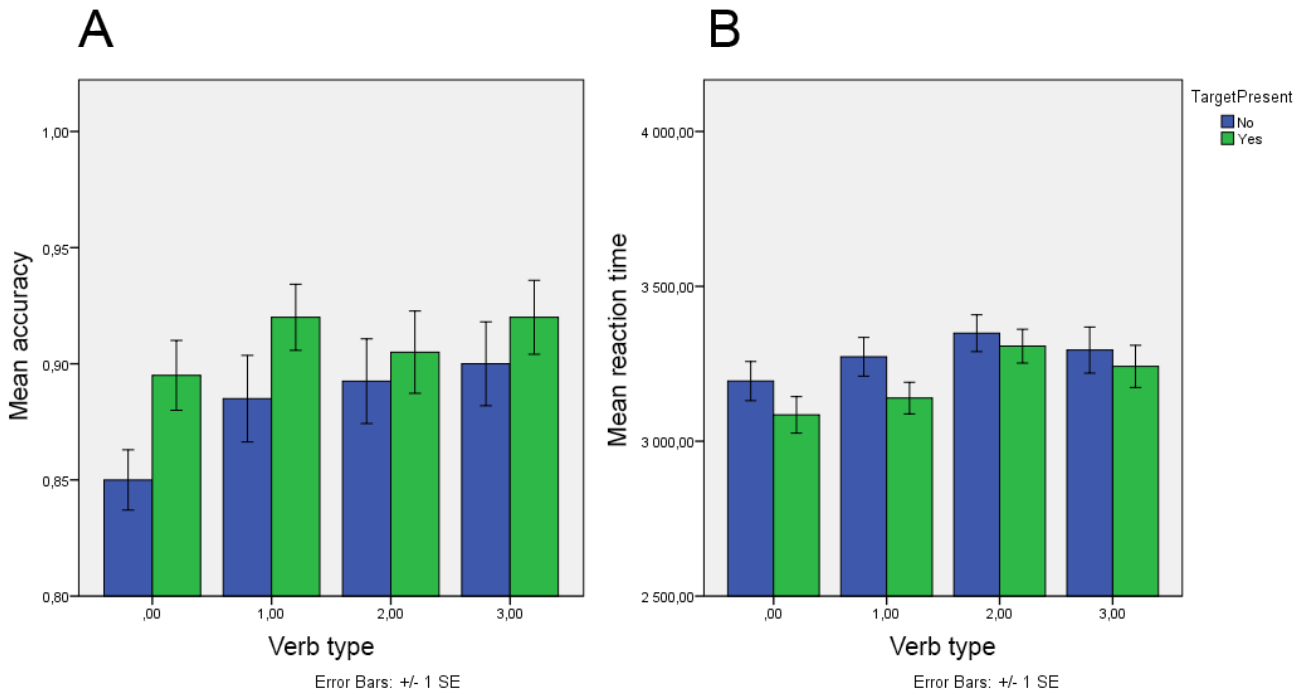


Figure 2: Mean accuracy and reaction time for each verb category and target presence condition.

The baseline category (where the verbs were not constraining) received the lowest accuracy scores of all the verb categories on average. A post-hoc pairwise comparison of verb categories found that this difference was significant when compared to categories 1 and 3, as seen in

Table ***2***.

Table 2: Pairwise comparison of accuracy across verb categories

Verb category	Compared against category	Mean difference	Std. error	Sig.	Effect size (Cohen's *d*)
0 (baseline)	1	-.030	.009	.016	-.067
	2	-.026	.012	.198	-.267
	3	-.038	.011	.017	-.928

Since the baseline verb category included one sentence for each of the target objects used in the experiment, and verbs did not constrain object selection, it could serve as a suitable category for investigating whether there were some object items that participants found particularly hard to identify. In the target-present condition sentences containing the following objects scored lowest on mean accuracy: *cattle* (.3600, std. deviation = .4899), *kettle* (.5600, std. deviation = .5066) and *mansion* (.6000, std. deviation = .5000). In the target-absent condition sentences containing the following objects scored lowest on mean accuracy: *flour bags* (.4800, std. deviation = .5099), *cattle* (.5200, std. deviation = .5099), *lemming* (.6000, std. deviation = .5000), *jumper cables* (.6800, std. deviation = .4761) and *lemon* (.6800, std. deviation = .4761).

4.3. Target-present condition

In the target-present condition, anticipatory looks towards the target item reflects an understanding of the verb and an ability to map that verb onto a likely referent (Brock et al., 2008, p. 899). Figure 3 shows the development of gaze proportion towards the target object from the onset of the verb.

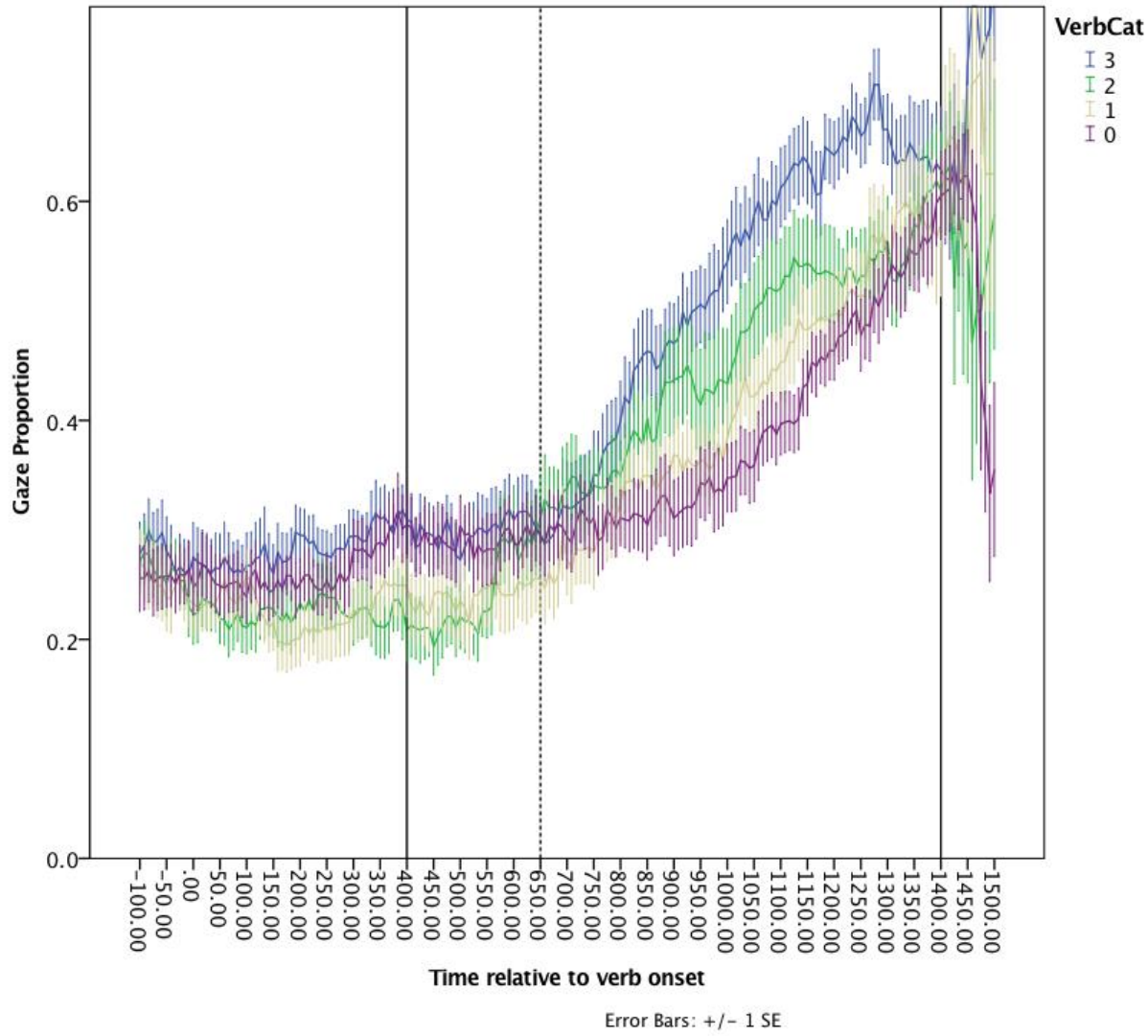


Figure 3: The X-axis shows gaze proportions for the target item in the target-present condition. The Y-axis shows time in milliseconds relative to verb onset. VerbCat 0 = baseline, 1 = least constraining, 2 = moderately constraining, 3 = most constraining. The solid vertical reference lines at 400 and 1400 ms denote the starting point and end point of the critical time window. The dashed vertical reference line denotes mean object onset time.

The mean target onset time was 651 ms post verb onset. Given the previously mentioned findings from Allopenna et al. (1998) that fixation on relevant objects start about 200 ms after the onset of its associated target word, we would not expect participants to start increasing their looks towards the target item before approximately 850 ms post verb onset if their gaze pattern was not mediated by the previous linguistic information (i.e. the verb).

Increasing gaze proportions for the target item is seen well before this point, with a larger effect for more constraining verbs than less constraining ones, and this larger effect for for more constrained verbs is seen to last until around 1400 ms post verb onset.

In order to compare the different verb categories, gaze proportions towards the different items were averaged over the course of a critical time window. This window was defined on the basis of visual inspection of the graph in Figure 3 in order to cover the time stretch when auditory stimuli mediated looks to the target object, and spanned from 400 ms to 1400 ms post verb onset. The resulting mean scores are presented in Table 3, and show that each level of constraint increased the probability of participants looking towards the target item in the critical time window.

Table 3: Proportion of looks towards the target item during the critical time window (400-1400 ms) for each verb category.

Verb category	Mean	Std. deviation
0 (baseline)	.37682	.098293
1 (least constraining)	.38316	.122983
2 (moderately constraining)	.41869	.150498
3 (most constraining)	.49112	.124457

A post-hoc pairwise comparison of verb categories was applied to determine whether the difference in gaze proportions across verb categories were significant. The results are presented in Table 4.

Table 4: Post-hoc pairwise comparison of gaze proportions across verb categories in target-present condition.

Verb category	Compared against category	Mean difference	Std. error	Sig.	Effect size (Cohen's *d*)
0 (baseline)	1	-.006	.020	1.000	-.067
	2	-.042	.033	1.000	-.261
	3	-.114	.025	<.001	-.928
1 (least constraining)	2	-.036	.036	1.000	-.196
	3	-.108	.023	<.001	-.659
2 (moderately constraining)	3	-.072	.038	.409	-.384

As seen in Table 4 the analysis found that gaze proportions for verb category 0 was not significantly different from category 1 or 2, and that results for category 1 was not significantly different from 2 either. However, gaze proportions for verb category 3 (the most constraining category) were significantly larger than proportions for both category 0 (the baseline category) and 1 (the least constraining of the test verb categories). There was no significant difference between categories 2 and 3. In other words, the only verb category that led to significantly increased looks to the target object compared to other categories was the most constraining one, and only for the least constraining of the three categories as well as the baseline category. The significance was adjusted for multiple comparisons using the Bonferroni correction.

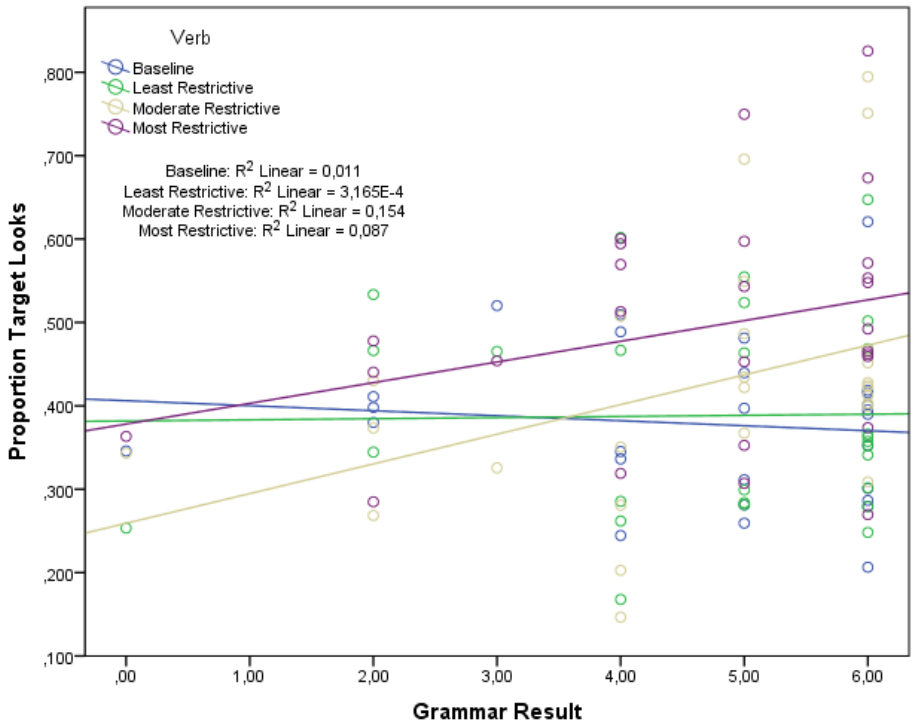


Figure 4: Relationship between gaze proportions for target-and results in the grammar test. Only for participants scoring C1 or C2 in the grammar test does the pattern of increased looks towards targets in verb categories 2 and 3 compared to 0 and 1 emerge.

An analysis of covariance (ANCOVA) was applied to the data in order to determine if the language measures (the grammar and vocabulary results) could predict the effect that the verb category had on gaze proportions in the critical time window. A test of within-subject effects showed a significant effect of grammar score, $F(3,66) = 3.202$, $p = .029$, $\eta^2 = .127$. However, it showed no significant effect of vocabulary score, $F(3,66) = 1.516$, $p = .218$, $\eta^2 = .064$. Only for participants with a sufficiently high grammar result (C1 or C2), did the effect of increased gazes for more constraining verb categories become visible. Figure 4 illustrates this. A test of

between-subjects effects showed no significant effect of grammar ($F(1,22) = .641$, $p = .432$, $\eta^2 = .028$) or vocabulary results ($F(1,22) = 1.118$, $p = .302$, $\eta^2 = .048$).

4.4. Target-absent condition

In the target-absent condition, the phonological effect of the cohort competitor was the main point of interest. Increased gaze proportions for the cohort competitor in trials where the target item was not present on the screen would indicate that the cohort competitors were effective, and that a phonological effect mediated gaze patterns in addition to the effect of verb semantics found in the target-present condition. In addition, if a decreased probability of looks towards cohort competitors in the more constraining verb categories compared to the less constraining ones was found, this could indicate that verbs whose argument structure requires very specific items in its object position discouraged participants from looking towards unlikely objects.

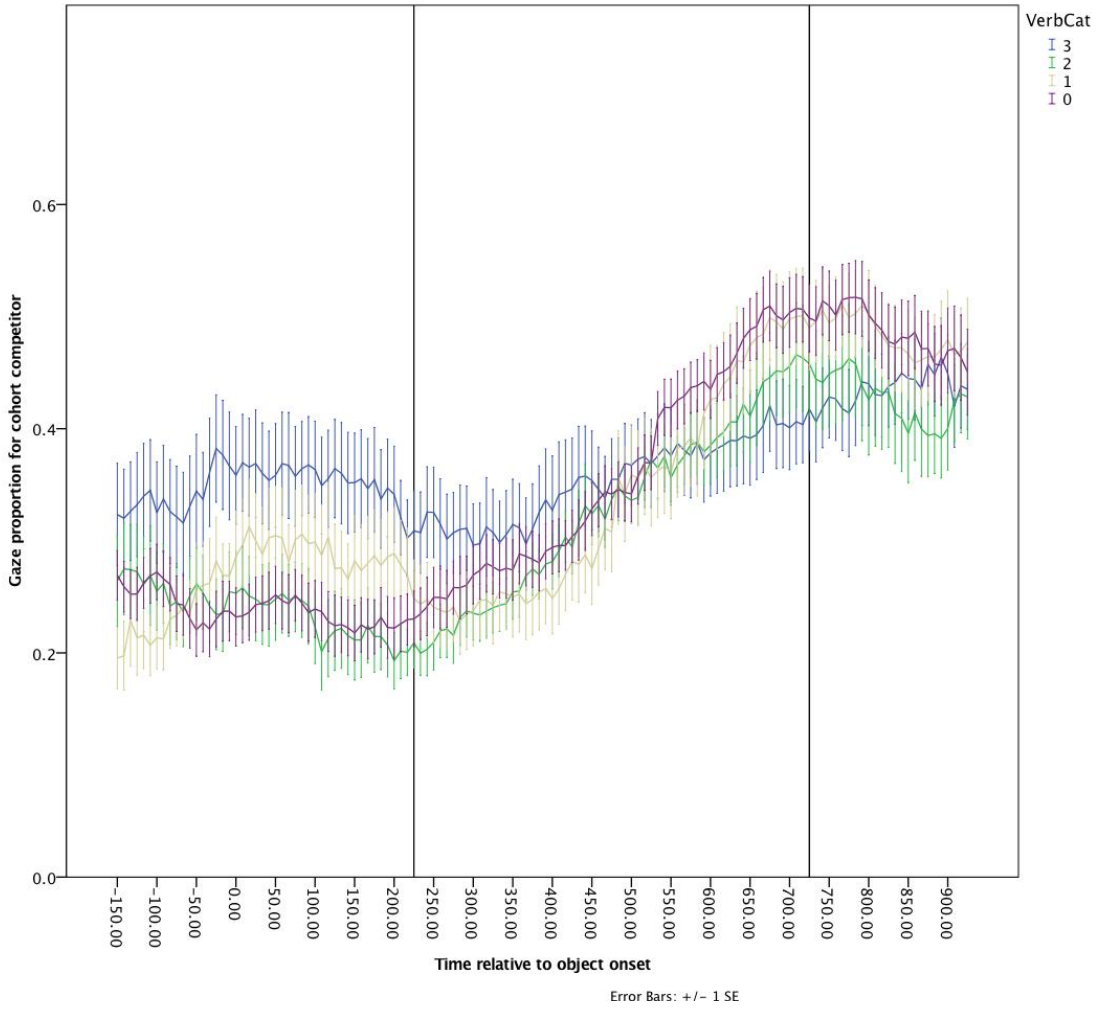


Figure 5: The X-axis shows gaze proportions for the competitor item in the target-absent condition. The Y-axis shows time in milliseconds relative to object onset. VerbCat 0 = baseline, 1 = least constraining, 2 = moderately constraining, 3 = most constraining. The solid vertical reference lines at 220 and 720 ms denote the starting point and end point of the critical time window.

Figure 5 shows the development of gaze proportion towards the cohort competitor object from the onset of the object. As in the target-present condition, gaze proportions towards the different items were averaged over the course of a critical time window in order to compare the different verb categories. This window was defined on the basis of visual inspection of the graph in Figure 5 in order to cover the time when auditory stimuli mediated looks to the competitor object, and spanned the time from 220 ms to 720 ms post object onset. The resulting mean scores are presented in Table 5, and show that the two most constraining verb categories have slightly lower scores than the least constraining and baseline categories.

Table 5: Proportion of looks towards the competitor item during the critical time window (220-720 ms) for each verb category.

Verb category	Mean	Std. deviation
0 (baseline)	.353	.031
1 (least constraining)	.357	.021
2 (moderately constraining)	.336	.027
3 (most constraining)	.326	.017

As in the target-present condition, a post-hoc pairwise comparison of verb categories was applied to determine whether the difference in gaze proportions across verb categories was significant. The comparison found no significant difference in gaze proportions for the different verb categories ($p > .05$). In other words, the constraint of the argument structure did not influence the probability of looks towards cohort competitors.

In order to check whether the phonological effect of the competitors was significantly different from chance level (a 25% chance of looking towards any of the four images), a T-test was applied to the gaze proportion scores. The test showed that gaze proportions for all verb categories were significantly above chance: the baseline category got the result $t(24) = 5.208$, $p < .001$; category 1 got the result $t(24) = 3.146$, $p = .004$; category 2 got the result $t(24) = 4.581$, $p < .001$; category 3 got the result $t(24) = 3.297$, $p = .003$. In other words, a phonological effect of cohort competitor was found for all verb categories.

5. Discussion

5.1. Gaze patterns in target-present condition

Altmann and Kamide (1999) were the first to find evidence from the visual world paradigm that the gaze pattern of native speakers is mediated by the information inherent in the argument structure of verbs, and that they make predictions of what linguistic information will follow. As such native speakers are shown to be very sensitive to argument structure. The results of the present study arguably showed a certain sensitivity to information inherent in verb meaning in second language users. When the target images were present on screen, the results show that the participant group as a whole looked significantly more towards the target image in the time from 400 to 1400 ms post verb onset if the sentence they heard contained the most constraining verbs compared to the baseline category and the least constrained verb category. This finding suggests a certain sensitivity to the argument structure of verbs in the participant group. It is, however, noteworthy that when the present data is compared to a native English speaking control group of a similar experiment carried out by Brock et al. (2008), the increase of target gaze proportions in constrained sentences becomes prominent almost 200 ms later in the present L2 user group (around 650 ms post verb), as the native speaker group increased looks at around 480 ms post verb (p. 899).

It must be noted that the majority of the "critical time window" that averages were calculated across takes place after the onset of the target object, and as such, I will not suggest that the average gaze proportion scores can be entirely attributed to participants predicting the target object based on their sensitivity to argument structure. However, as stated before, previous work within the visual world paradigm have shown a 200 ms latency between the onset of auditory stimuli and increased looks towards the visual representation associated with this stimuli (Allopenna et al., 1998, p. 428). This means that if participants had only heard the object instead of the whole sentence, increased looks towards the target image should not have been visible until around 200 ms after target-onset, which on average corresponds to the 850 ms mark on the X-axis of Figure 3 in the results section. When observing that target gaze proportions at this point had already increased substantially, I will suggest that participants were able while processing the auditory input to employ their lexical knowledge of the previously occurring verbs in a predictive manner to expect what visual item was the most probable object of the verb. The fact that each level of constraint in verb category showed increased gaze proportions compared to the previous one at this point further strengthens the idea that the gaze pattern was mediated by information found in the verb at

this point, even if the effects were not significant in all cases when comparing each verb constraint level to the previous one.

Another central finding is that this pattern of increased target gazes for constraining verbs depends on a sufficiently high score in the grammar tests. To the extent that the grammar test used in the present study could reliably predict the participants' actual CEF proficiency levels, only participants at the levels of C1 and C2 showed this gaze pattern. Participants scoring below C1 did not show any coherent pattern with regards to gaze proportions for different verb categories, which could suggest that this is the proficiency level at which sensitivity to argument structure is developed enough to employ it in a predictive manner in language processing. It must be noted here that the group on average scored rather high on the grammar test, with only 9 of the 25 participants scoring lower than C1, so any conclusions drawn for the lower grammar results participants are based on fairly scarce data.

5.2. Gaze patterns in target-absent condition

The results showed a significant phonological effect in the target-absent condition, with looks towards the cohort competitor increasing from around 220 ms post target word onset. This finding adheres closely to the findings of Allopenna et al. (1998) who observed increased gaze proportions for cohort competitor items from around 200 ms post target onset (p. 428). It is not surprising that the timing of the phonological effect is close to findings from native speaker studies, as the effect does not rely on grammar aptitude, only lexical knowledge of the items referenced by the images displayed on the screen.

Figure 5 presented in the results section showed that at the peak of the phonological effect, which was 720 ms post target, the unconstrained baseline category displayed the strongest phonological effect, while the most constrained category displayed the weakest effect. This can be read to indicate that the argument structure of the previously occurring verbs influenced gaze patterns and deterred looks towards unlikely candidates for the sentence object. However, when averaged across the critical time window, the difference between verb categories were not found to be significant, and the baseline category even had lower gaze proportions than verb category 1. In the previously mentioned experiment carried out by Brock et al. (2008), which used a similar experimental design, native speaker data showed that the phonological effect of the competitor in a target-absent condition was significantly weakened by constraining verbs compared to their baseline condition, which in their case used the verb *choose* for all sentences (p. 899). Their study used only two verb categories – neutral and constrained, but seeing as the present study found no significant difference for any

two categories, the point remains the same: the second language users show a weaker sensitivity to argument structure compared to native speakers in the case when only the competitor and unrelated distractors are present on the screen. This could indicate that the phonological effect is stronger compared to the semantic effect in the language processing of the present L2 user group than in native speakers.

5.3. Accuracy and reaction time

Overall, the results showed high mean accuracy scores across all verb categories. However, certain individual items received relatively low accuracy scores. This indicates that there are some issues with these items that could also have influenced the gaze patterns and made the findings of the experiment less obvious than if all items had very high accuracy scores. The sentences containing the target items "cattle", "kettle" and "mansion" received the lowest mean accuracy scores in the target-present condition of the baseline sentence category. In the target-absent condition, "flour bags", "cattle", "lemming" and "lemon" show the lowest accuracy rates in the same sentence category. These sentences all contained the non-constraining verbs "chose", "picked", "selected" and "took", such that verbal semantics should not influence participants' answer. In the target-present condition, wrong answers could indicate that the participants are not familiar with the words, and that they associated different nouns with the images. In hindsight, most of the stimulus words that received the lowest accuracy scores were probably too uncommon for most intermediate L2 users to know them, and a way of remedying this problem would have been to check the proposed target items against a corpus beforehand to determine their prevalence in use, which was not done. It could be that participants took the target image for "cattle" to depict "cows", the image for "mansion" to depict "house" and the image for "lemming" to depict some other rodent. In the target absent condition, an additional possible explanation is that the presence of the cohort competitor prompted an incorrect "yes" answer from the participants. If this was the case, it affirms the phonological effect of the competitor, and as such is of value to the study.

The decision was made to keep all items in the analysis despite low accuracy scores for certain items, given that the analysis still yielded significant results for effects of the verb categories on gaze proportions. Unfamiliarity with the particular words does not necessarily compromise their status as target objects when examining how verbal semantics influence gaze patterns. The verb "milk" will presumably mediate looks towards the image depicting "cattle" regardless of whether participants are familiar with this particular lexical item or not. However, these items are presumably not very effective as competitor items. The

phonological similarity of "cattle" and "kettle" is irrelevant if the image depicting cattle prompts activation of the lexical item "cows". Likewise, the identical onset syllables of "man" and "mansion" is irrelevant if participants activate the lexical item "house" when looking at the image depicting a mansion. It is possible that this issue affected the data. The issue was predicted when designing the experiment, which is the reason why "problematic" items were displayed as targets before they were displayed as competitors, in an effort to train participants to hold the connection between the auditory input (e.g. the word "cattle") and the visual input (e.g. the image of cows). It is however quite possible that this pre-emptive measure was insufficient to eliminate the problem. This issue highlights a potential weakness in all experimental designs that rely on visual stimuli functioning as competitors. The effectiveness of the competitor rests on an assumption that the image on the screen unambiguously activates a certain linguistic item in the participant's mind, which is not necessarily always the case.

Reaction times show that the participants on average reacted quicker in the target-present condition than in the target-absent one. This can be read to indicate that visual context influences and constrains language processing in L2 users. As such the results comply with the notion from Huettig et al. (2010) about bias towards greater processing speeds because of a very limited number of alternatives to choose from in the visual context.

5.4. Verb categorisation

The categorising of verbs was crucial to the design of the present thesis and the research question underlying the design. However, the actual measures for grouping different verbs into different categories were formulated in a rather arbitrary fashion. It was decided beforehand that the study was to include three levels of constraint. Among the predefined verbs used in the native speaker survey discussed in section 3.4.1, the desire was to devise a formula that would divide these verbs relatively equally into three different categories. The rule of having 55% of all answers as a lower limit, and requiring less than seven or less than four different answers to constitute 55% in order for the verb to be put in the moderately or most constraining categories respectively, were created from trial and error with the data. These grouping measures were devised for simplicity, and paid no attention to the semantics of the answers provided by the native speaker participants. As such, it was entirely possible for answers that were semantically similar, for instance hyponym-hyperonym relations like "jacket" and "clothes" (for the verb *remove*) or near-synonyms like "fire" and "flames" (for the verb *extinguish*), to negatively affect the "constraint level" of the verbs of the study just as

much as answers that are clearly much less semantically related like “gun” and “car” (for the verb *load*). This is a potential weakness in the design, which was non the less unheeded for the sake of simplicity and due to time constraints.

Looking at the results, it seems that the weakness of the verb categorisation could be somewhat reflected in the gaze data. As was seen in the results, neither the least constrained nor the moderately constrained categories yielded significantly more looks than the baseline category in the critical time period, although it should be noted that the target gaze proportions for the target-present condition did line up exactly like expected in the time window from 800-1300 ms post verb onset. This latter finding suggests that while there was a weakness in the method of verb categorisation, this weakness was not completely decisive to the outcome, and the gaze behaviour of participants partly matched the expectations. It is possible that a different categorisation could have yielded clearer answers to the research question treated in this thesis. Fewer verb categories, for instance two levels of constraint in addition to the baseline could possibly have resulted in a larger difference in the pairwise comparisons across verb categories. I believe it is a strength of the method that it takes into account common usage of verbs, by eliciting “first whim” answers from native speakers, and as such is sensitive to the fact that certain verbs have highly conventionalised uses, which L2 users are also likely to know and expect when processing the verbs. It would, however, possibly enhance the results if a different method for categorisation which takes into account the semantic properties of the answers from native speakers was employed. One possibility for later studies is to quantify and take into account how many of the objects on the screen are semantically feasible candidates for object of that verb, and employ this information when categorising the verbs. There were several instances in the present study where more than one image represented a semantically acceptable object of the trial verb, but this fact was not considered in the experimental design or in the analysis.

6. Conclusion

The aim of this study was to investigate how sensitive Norwegian speakers of English were to the argument structure of English verbs, and how this trait is developed. As mentioned in the introduction, the present thesis only treats the data from one of the two groups that were tested. As such, the data only paints half the picture, and a comparison across groups would provide a richer insight into the development of second language argument structure, seen in light of differences in maturity and language proficiency. However, certain conclusions can be drawn on the basis of the data treated in the present thesis.

As expected at the onset of the study, the findings suggested a certain sensitivity to argument structure in the younger group in the form of increased gaze proportions for the most constrained verb category compared to all other categories. Contrary to the expectations, no significantly different gaze proportions were found for the other verb categories. This lack of difference for the other verb categories could be a result of poor verb categorisation during the construction of the stimuli. Fewer categories and other measures for categorisation could possibly have provided clearer differences across categories. A comparison between gaze proportions and grammar results suggested that only participants with a high English grammar proficiency were able to use the information in the argument structure to look more towards appropriate objects given constrained verbs compared to less constrained ones, which could suggest that a high level of grammar aptitude is necessary to use argument structure of verbs in a predictive manner in language processing.

The data from the target-absent condition showed a certain phonological effect in the form of increased looks towards the cohort competitor item. However, contrary to expectations, the constraint of the verb did not influence competitor looks, which could suggest that the phonological effect is stronger than the effect of argument structure in the target-absent condition for L2 users at this level of language aptitude.

7. References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*(4), 419-439. doi:10.1006/jmla.1997.2558

Altmann, G. T. M. (1997). Words and how we (eventually) find them. *The ascent of Babel: an exploration of language, mind, and understanding* (pp. 1-32). Oxford: Oxford University Press.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition: International Journal of Cognitive Science, 73*(3), 247-264. doi:10.1016/S0010-0277(99)00059-1

Brock, J., Norbury, C., Einav, S., & Nation, K. (2008). Do individuals with autism process words in context? Evidence from language-mediated eye movements. *Cognition: International Journal of Cognitive Science, 108*(3), 896-904. doi:10.1016/j.cognition.2008.06.007

Brooks, P. J., & Tomasello, M. (1999). How children constrain their argument structure constructions. *Language, 75*(4), 720-738. doi:10.2307/417731

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6*(1), 84-107. doi:http://dx.doi.org/10.1016/0010-0285(74)90005-X

Dekydtspotter, L., Schwartz, B. D., Sprouse, R. A., & Bullock, G. (2008). *Locative verbs in Korean-English and the universal grammar question revisited: Disentangling lexical semantics, syntax and pragmatics*. Ms., Indiana University.

Dimitrova-Vulchanova, M., & Dekova, R. (2007). Event lexicalization across languages. *Southern Journal of Linguistics, 29*, 36-51.

Dimitrova-Vulchanova, M., & Weisgerber, M. (2007). Encoding and relevant bits of information. *Southern Journal of Linguistics, 29*, 52-76.

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

Huettig, F., Rommers, J., & Meyer, A. S. (2010). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137*(2), 151-171. doi:10.1016/j.actpsy.2010.11.003

Kamide, Y., Altmann, G. T. M., & 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. doi:10.1016/S0749-596X(03)00023-8

Koenig, J.-P., Mauner, G., & Bienvenue, B. (2003). Arguments for adjuncts. *Cognition, 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.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, Mass: MIT Press.

Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, Mass: MIT Press.

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

Schwartz, B. D., Dekydtspotter, L., & Sprouse, R. A. (2003). *Pouring fire with gasoline: Questioning conclusions on L2 argument structure*. Paper presented at the Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference (GASLA 2002).

Stringer, D. (2010). The gloss trap. In Z. H. Han (Ed.), *Linguistic relativity in SLA* (pp. 102-124). Bristol: Multilingual Matters.

Treffers-Daller, J., & Calude, A. (2015). The role of statistical learning in 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

Appendix A: Consent form for L2 experiment

Forespørsel om deltakelse i forskningsprosjektet
"Andrespråksbrukeres behandling av engelske verb"

Bakgrunn og formål
Vi er to masterstudenter ved Institutt for språk og litteratur ved NTNU som jobber med et forskningsprosjekt. Vi trenger deltakere både fra videregående skole og universitetet.
Målet med prosjektet vårt er å undersøke hvordan norske fremmedspråksbrukere av engelsk prosesserer (hvordan hjernen behandler) engelske verb. Deltakelse er frivillig, og vi er svært takknemlige for alle som har mulighet til å være med.

Hva innebærer deltakelse i studien?
Deltakerne vil gjennomføre tre elektroniske spørreskjemaer/tester på PC, samt en test hvor deltakerne sitter foran en datamaskin og får høre opptak av engelske setninger samtidig som de ser ulike bilder på skjermen. I denne testen bruker vi en "eye-tracker" (et kamera som kun registrerer øyebevegelse og lagrer informasjon om hvor på skjermen man ser). Testene vil samlet sett ta omtrent en time.

Hva skjer med informasjonen om deg?
Alle personopplysninger vil bli behandlet konfidensielt. En "koblingsnøkkel" (et deltakernummer) vil knytte navnene til resultatene. For skoleelevene vil det kun være læreren som har tilgang til listen som knytter navnene til deltakernummeret, og denne skal lagres utilgjengelig for uvedkommende. Enkeltpersoner vil ikke kunne gjenkjennes i publikasjonen.

Prosjektet skal etter planen avsluttes juni 2016. Personopplysninger vil da slettes, slik at datamaterialet er anonymisert

Frivillig deltakelse
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:

Roger Johnsen, masterstudent ved Institutt for språk og litteratur, NTNU
Tlf.: 416 75 224
E-post: roger.johnsen88@gmail.com

Anders Schärer Reine, masterstudent ved Institutt for språk og litteratur, NTNU
E-post: andersscharer@gmail.com

Mila Vulchanova, professor ved Institutt for språk og litteratur, NTNU
E-post: mila.vulchanova@ntnu.no

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

Samtykke til deltakelse i studien
Jeg har mottatt informasjon om studien, og er villig til å delta:

(Signert av prosjektdeltaker, dato)

Appendix B: Background questionnaire for L2 participants

Bakgrunnsinformasjon for forskningsprosjekt om andrespråksforståelse

Tusen takk for at du har sagt ja til å delta i vårt forskningsprosjekt om andrespråksforståelse. I dette skjemaet ber vi om bakgrunnsinformasjon som er nødvendig for at resultatene fra undersøkelsen skal kunne brukes.

Informasjonen som du oppgir vil bli behandlet uten direkte gjenkjennende opplysningser. En kode knytter deg til dine opplysninger gjennom en deltakerliste. Det er kun autorisert personell knyttet til prosjektet som har adgang til deltakerlisten og som kan finne tilbake til informasjonen. Del B, C og D av dette skjemaet vil bare oppbevares med koden. All informasjon vil bli anonymisert ved prosjektslutt. Det vil ikke være mulig å identifisere deg i resultatene av studien når disse publiseres.

Vi ber deg legge merke til at skjemaet har totalt 7 sider.

Roger Johnsen / Anders Schärer Reine
Studenter ved lektorutdanningen med master i språk, NTNU

Deltakerkode:

Del A: Personlig informasjon

Studieretning og trinn:_____

Fødselsår_____

Kjønn □ **Kvinne** □ **Mann**

Bostedskommune_____

Deltakerkode:

Del B: Språklig bakgrunn

Morsmål

Er norsk morsmålet ditt?

□ Ja □ Nei

Hvis ja, har du andre morsmål i tillegg?

□ ja □ Nei

Hvis ja, hvilke(t) språk? _____

Hvilket språk bruker dere hjemme?_____

Hvor ofte leser du tekst skrevet på norsk?

□ Hver dag □ Flere ganger i uka □ Et par ganger i uka □ Av og til □ Aldri

Hvor ofte skriver du tekst på norsk?

□ Hver dag □ Flere ganger i uka □ Et par ganger i uka □ Av og til □ Aldri

Engelsk og andre fremmedspråk

I <u>engelsk,</u> hvordan vurderer du ferdighetene dine på hvert av disse områdene?

	Grunnleggende	**Middels**	**Avansert**	**Flytende**
Lesing				
Skriving				
Snakke				
Lytte				
Totalt				

Har du bodd i, eller hatt lengre opphold i, et land hvor engelsk er hovedspråk?

□ Ja □ Nei

Hvis ja, hvor lenge varte oppholdet/oppholdene?_____

Har du vært på kortere (under 14 dager) reise i et land hvor engelsk er hovedspråk?

□ Ja □ Nei

Har du bodd i, eller hatt lengre opphold i, et land hvor annet enn engelsk er hovedspråk?

□ Ja □ Nei

Hvis ja, hvor var det, og hvor lenge varte oppholdet/oppholdene?_____

_____

Hvilke språk kan du utover morsmålet ditt og engelsk?

Språk	**Nivå**			
	Grunnleggende	**Middels**	**Avansert**	**Flytende**
Tysk				
Fransk				
Spansk				
-Angi språk				
-Angi språk				
-Angi språk				

Hvor ofte leser du tekster på engelsk?

□ Hver dag □ Flere ganger i uka □ Et par ganger i uka □ Av og til □ Aldri

Hvor ofte skriver du tekster på engelsk?

□ Hver dag □ Flere ganger i uka □ Et par ganger i uka □ Av og til □ Aldri

Hvor ofte lytter du til/hører du engelsk?

□ Hver dag □ Flere ganger i uka □ Et par ganger i uka □ Av og til □ Aldri

Hvor ofte ser du på engelskspråklige serier/filmer?

□ Hver dag □ Flere ganger i uka □ Et par ganger i uka □ Av og til □ Aldri

Når du ser på engelskspråklige filmer, hvilken av disse alternativene bruker du mest?

□ Undertekst på norsk □ Undertekst på engelsk □ Ingen undertekst

Hvor ofte ser du på engelskspråklige tegneseriefilmer/serier?

□ Hver dag □ Flere ganger i uka □ Et par ganger i uka □ Av og til □ Aldri

Hvor ofte spiller du engelskspråklige data/TV-spill?

□ Hver dag □ Flere ganger i uka □ Et par ganger i uka □ Av og til □ Aldri

Hvilke type spill spiller du? _____

Hvor mange timer cirka per dag? _____

Hvor mye TV ser du på hver dag?

□ 7 timer eller mer □ 5-6 timer □ 3-4 timer □ 1-2 timer □ aldri eller nesten aldri

Del C: Andre faktorer i språklæring

Har du, eller har du hatt, problemer med synet utover normal brillebruk?

□ Ja □ Nei

Har du, eller har du hatt, problemer med hørselen?

□ Ja □ Nei

Har du, eller har du hatt, språkvansker av noe slag (spesifikke språkvansker, lese-/lærevansker eller lignende)?

□ Ja □ Nei

Har du, eller har du hatt, andre diagnoser som kan tenkes å påvirke språklæring (ADHD, autisme eller lignende)?

□ Ja □ Nei

Er du venstrehendt?

□ Ja □ Nei

Del D: Vokabulartest og grammatikktest

Resultat vokabulartest:

Resultat grammatikktest:

Appendix C: Information sheet for L1 survey

Norwegian University of Science and Technology

Sensitivity to argument structure in second language

Participant Information Sheet

Thank you for your interest in this study. Before you decide whether to take part, please read the following information carefully (this sheet is for you to keep).

What is this research looking at?
As part of research into the linguistic capabilities of second language English speakers, we need to know how native English speakers would use some common English verbs in regular sentences. This examination is part of the initial stages of the research.

Do I have to take part?
It is up to you to decide to join the study. We will describe the study in this information sheet. If you agree to take part, you will be presented with an online questionnaire which you fill out anonymously, and by submitting this questionnaire, you consent to participation. You are free to withdraw at any time, without giving a reason, and if you do not submit the questionnaire, your answers will not be saved in any database. Withdrawing would not affect you in any way.

What will happen if I agree to take part?
You will have to fill out an online questionnaire, which will take about 15-20 minutes to complete. At first we ask for some basic information about you (age, gender, number of languages spoken at home, education level). Then you will be asked to complete a number of incomplete sentences by writing short answers in text boxes. Participants will either receive course credit or a token reward of 3 pounds for their participation.

Are there any problems with taking part?
There are no problems or disadvantages to taking part in this study.

Will it help me if I take part?
No.

How will you store the information that I give you?
All information which you provide during the study will be stored in accordance with the 1998 Data Protection Act and kept strictly confidential. The chief investigator will be the custodian of the anonymous research data. No identifiable data will be collected. All anonymized results will be stored indefinitely on a password protected computer. Only the research team will have access to the data.

How will the data be used?
The data will be part of a research project at the Norwegian University of Science and Technology and might be presented in journals and conferences. Participants will not be identifiable in the presented data.

What happens if I agree to take part, but change my mind later?
Once the data is submitted, it is no longer possible to withdraw from the survey. However, you may choose to withdraw from the survey at any point while filling out the questionnaire, and the data will not be saved in any database until you click to submit it at the end.

How do I know that this research is safe for me to take part in?
All research in the University is looked at by an independent group of people, called a Research Ethics Committee, to protect your safety, rights, wellbeing and dignity.

You are under no obligation to agree to take part in this research.
If you do agree you can **withdraw at any time without giving a reason.**

Reseacher Contact details:
Roger Johnsen
rogej@stud.ntnu.no
Tel: +47 41675224

Supervisor Contact details:
Mila Vulchanova
mila.vulchanova@ntnu.no
Tel: +47 73596791

Do also contact us if you have any worries or concerns about this research.

Appendix D: Consent form for L1 survey

Norwegian University of Science and Technology

Consent Form

"Sensitivity to argument structure in second language"

Name of Researcher: Roger Johnsen/Anders Schärer Reine

Please initial all boxes

1. I have read and understand the information sheet (Sensitivity to argument structure in second language) and have had the opportunity to questions and have had these answered satisfactorily. ☐
2. My participation is voluntary and I know that I am free to withdraw at any time, without giving any reason and without it affecting me at all ☐
3. I know that no personal information (such as my name) will be shared outside of the research team or published in the final report(s) from this research ☐
4. I agree to take part in the above study ☐

Participant's signature…….. Date……..

Reseacher Contact details:
Roger Johnsen
rogej@stud.ntnu.no
Tel: +47 41675224

Supervisor Contact details:
Mila Vulchanova
mila.vulchanova@ntnu.no
Tel: +47 73596791

Appendix E: Table listing stimuli sentences and images

Stimulus sentence	Target image	Replacement *(in target-absent condition)*	Competitor	Distractor 1	Distractor 2
Least constrained verb category					
Alex memorised the notes carefully.	notes	light bulb	nose	laces	laser gun
Mary played the bagpipe happily.	bag pipe	back pack	bags	mansion	man
Sam guarded the bell carefully	bell	jumper	belt	flower	flour bags
Charlie removed the jumper cables gently.	jumper cables	belt	jumper	bacon	baby
Alex rubbed his nose regularly.	nose	lifeboat	notes	laces	laser gun
Mary watched the bags carefully.	bags	back door	bag pipe	mansion	man
Sam wore the back pack happily.	back pack	bagpipe	back door	cattle	kettle
Charlie bought the laser gun happily.	laser gun	lemon	laces	notes	nose
Mary broke the bagpipe accidentally.	bagpipe	back pack	bags	mansion	man
Jesse carried the jumper cables regularly.	jumper cables	belt	jumper	bacon	baby
Sam changed the light bulb quickly.	light bulb	notes	lifeboat	lemming	lemon
Charlie cut the lemon carefully.	lemon	laser gun	lemming	light bulb	lifeboat
Alex dropped the kettle accidentally.	kettle	man	cattle	backpack	back door
Jesse examined the bell carefully.	bell	jumper	belt	flowers	flour bags
Sam filled the back pack quickly.	pack	bagpipe	back door	cattle	kettle
Mary judged the man quickly.	man	kettle	mansion	bags	bagpipe
Moderately constrained verb category					
Alex loaded the lifeboat quickly.	lifeboat	nose	light bulb	lemming	lemon
Jesse mended the belt carefully.	belt	jumper cables	bell	flowers	flour bags
Sam missed the lifeboat accidentally.	lifeboat	nose	light bulb	lemming	lemon
Charlie ordered the flowers eagerly.	flowers	bacon	flour bags	bell	belt
Mary painted the mansion carefully.	mansion	cattle	man	bags	bagpipe
Jesse pinched his nose regularly.	nose	lifeboat	notes	laces	laser gun
Sam pushed the back door gently.	back door	bags	backpack	cattle	kettle
Charlie served the bacon eagerly.	bacon	flowers	baby	jumper	jumper cables
Alex typed the notes hastily.	notes	light bulb	nose	laces	laser gun
Mary visited the mansion regularly.	mansion	cattle	man	bags	bagpipe
Sam delivered the flour bags quickly.	flour bags	baby	flowers	bell	belt
Charlie entertained the baby happily.	baby	flour bags	bacon	jumper	jumper cables
Alex frightened the cattle accidentally.	cattle	mansion	kettle	backpack	back door
Jesse hunted the lemming eagerly.	lemming	laces	lemon	light bulb	lifeboat
Sam impressed the baby happily.	baby	flour bags	bacon	jumper	jumper cables
Mary collected the bags quickly.	bags	backdoor	bagpipe	mansion	man
Most constrained verb category					
Alex milked the cattle carefully.	cattle	mansion	kettle	backpack	back door
Jesse squeezed the lemon gently.	lemon	laser gun	lemming	light bulb	lifeboat
Sam stroked the lemming carefully.	lemming	laces	lemon	light bulb	lifeboat
Charlie tied the laces quickly.	laces	lemming	laser gun	notes	nose
Alex tightened the laces carefully.	laces	lemming	laser gun	notes	nose
Jesse fastened the belt eagerly.	belt	jumper cables	bell	flowers	flour bags
Sam locked the back door accidentally.	back door	bags	backpack	cattle	kettle

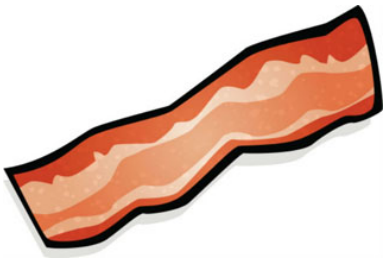
Stimulus sentence	Target image	Replacement *(in target-absent condition)*	Competitor	Distractor 1	Distractor 2
Charlie watered the flowers regularly.	flowers	bacon	flour bags	bell	belt
Mary arrested the man quickly.	man	kettle	mansion	bags	bagpipe
Jesse boiled the kettle hastily.	kettle	man	cattle	backpack	back door
Sam emptied the flour bags accidentally.	flour bags	baby	flowers	bell	belt
Charlie fired the laser gun quickly.	laser gun	lemon	laces	notes	nose
Alex fried the bacon hastily.	bacon	flowers	baby	jumper	jumper cables
Jesse ironed the jumper carefully.	jumper	bell	jumper cables	bacon	baby
Sam knitted the jumper carefully.	jumper	bell	jumper cables	bacon	baby
Charlie lit the light bulb quickly.	light bulb	notes	lifeboat	lemming	lemon
Baseline category					
Alex chose the jumper eagerly.	jumper	bell	jumper cables	bacon	baby
Jesse chose the notes regularly.	notes	light bulb	nose	laces	laser gun
Mary chose the bags quickly.	bags	back door	bagpipe	mansion	man
Charlie chose the man quickly.	man	kettle	mansion	bags	bagpipe
Alex chose the flowers eagerly.	flowers	bacon	flowers	bell	belt
Jesse chose the bacon quickly.	bacon	flowers	baby	jumper	jumper cables
Sam picked the light bulb eagerly.	light bulb	notes	lifeboat	lemming	lemon
Mary picked the mansion regularly.	mansion	cattle	man	bags	bagpipe
Alex picked the back door quickly.	back door	bags	backpack	cattle	kettle
Jesse picked the cattle eagerly.	cattle	mansion	kettle	backpack	back door
Sam picked the lemming happily.	lemming	laces	lemon	light bulb	lifeboat
Charlie picked the laces quickly.	laces	lemming	laser gun	notes	nose
Alex selected the jumper cables quickly.	jumper cables	belt	jumper	bacon	baby
Jesse selected his nose quickly.	nose	lifeboat	nose	laces	laser gun
Mary selected the bagpipe accidentally.	bagpipe	backpack	bags	mansion	man
Charlie selected the belt carefully.	belt	jumper cables	bell	flowers	flour bags
Alex selected the flour bags quickly.	flour bags	baby	flowers	bell	belt
Jesse selected the baby regularly.	baby	flour bags	bacon	jumper	jumper cables
Sam took the lifeboat accidentally.	lifeboat	nose	light bulb	lemming	lemon
Charlie took the bell eagerly.	bell	jumper	belt	flowers	flour bags
Alex took the backpack quickly.	backpack	bagpipe	back door	cattle	kettle
Mary took the kettle gently.	kettle	man	cattle	backpack	back door
Sam took the lemon gently.	lemon	laser gun	lemming	light bulb	lifeboat
Charlie took the laser gun regularly.	laser gun	lemon	laces	notes	nose

Appendix F: Overview of stimuli images

baby

back door

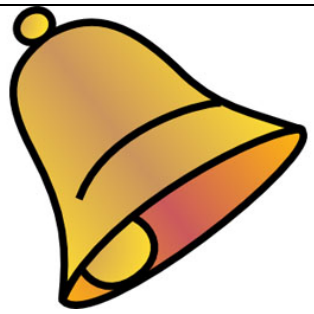
back pack



bacon

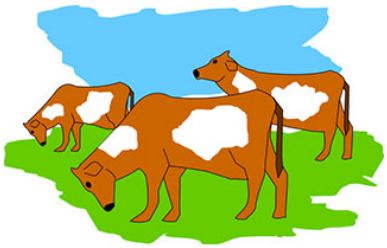
bagpipe

bags



bell

belt

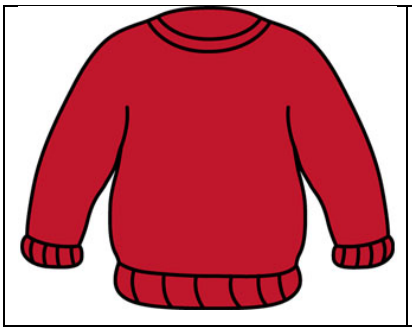


cattle

flour bags

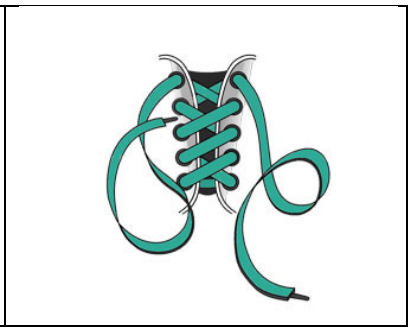
flowers

jumper cables

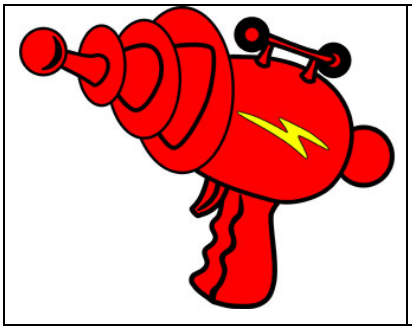


jumper

kettle

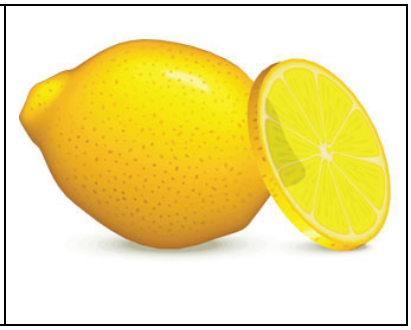


laces

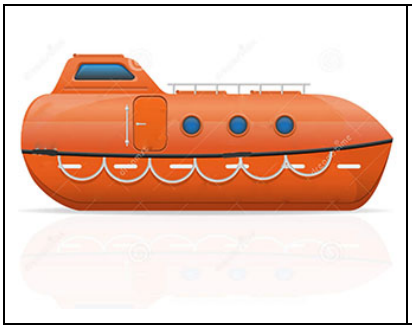


laser gun

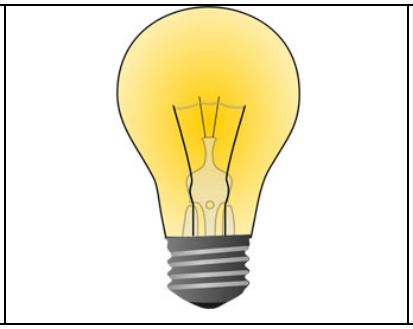
lemming



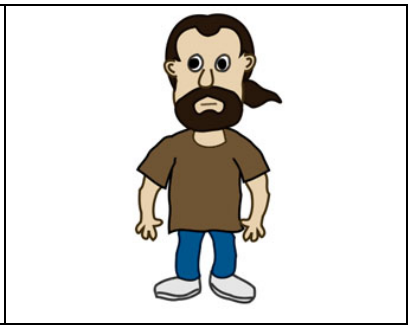
lemon



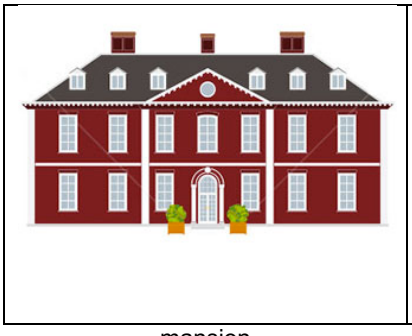
lifeboat



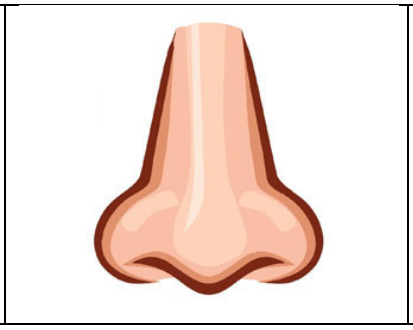
light bulb



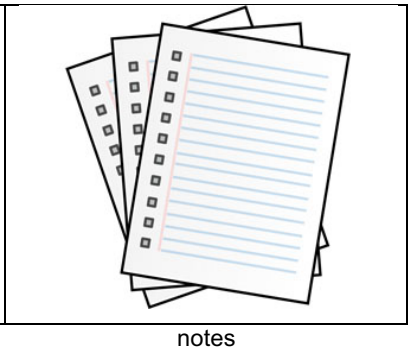
man



mansion



nose



notes