Master's thesis

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

Erlend Ness Rannem

Sensitivity to Second Language Argument Structure of Verbs

An Empirical Study with Norwegian Learners of English

Master's thesis in English Linguistics and Language Acquisition

May 2021

NTNU
Norwegian University of Science and Technology

Erlend Ness Rannem

Sensitivity to Second Language Argument Structure of Verbs

An Empirical Study with Norwegian Learners of English

Master's thesis in English Linguistics and Language Acquisition
May 2021

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

NTNU
Norwegian University of Science and Technology

Abstract

Language users continuously predict upcoming information in their L1 by exploiting information encoded in verbs' lexical entries. As such, native speakers are sensitive to the argument structure of verbs. But what happens when the linguistic context is shifted to a non-native language? This thesis tests whether and to what extent predictive language processing occurs in an L2, and investigates how the factors of verb constraint and verb frequency influence in this regard. 54 native Norwegian speakers of English responded to two language tests. One sentence completion test where sentences were grouped by different levels of constraint (most-, moderately-, and least constrained verbs) and frequency (high- and low frequency verbs); and one vocabulary test. The findings suggest that every participant displayed some level of sensitivity, irrespective of one's proficiency; instead, the relationship between proficiency and sensitivity held only a deterministic role in the most constraining contexts. In addition, the following constraint- and frequency effects were observed: more target response overlaps and less unique responses were elicited the more constraining a verb was (most > moderately > least) and the less frequent a verb was (low > high).

Acknowledgments

I realize, as I'm making the finishing touches on this project, that I soon will have to leave the stressful and frustrating yet enlightening and gripping bubble of writing your master's thesis behind. It's a bittersweet feeling to finally be finished considering the countless hours poured into it. The result is something I'm proud of, and I like to think the same description J.R.R. Tolkien's fantasy character *Treebeard* in *The Lord of the Rings* gives about his lovely language *Old Entish* applies to the process of writing this thesis as a whole: "It is a lovely language, but it takes a very long time to say anything in it, because we do not say anything in it, unless it is worth taking a long time to say, and to listen to" (2005, p. 465). Still, the thesis would not have been produced without the help of some astounding people.

Foremost, I'd like to express a sincere gratitude to my research supervisor Prof. Mila Vulchanova who has led me along the way; providing me with guidance and feedback during the development of this project. Her contribution to theoretical and practical insights has elevated my understanding of linguistics.

In addition, I'd also like to thank all my informants who volunteered to participate in my investigation. I'm appreciative each and every one of you sacrificed more or less 60 minutes of your day to reply to my questionnaire.

At last, but no less important, thanks to my partner and girlfriend, Tiril, who kept my spirits high these past 12 months. You have backed me up throughout the writing process and kept my sanity in check.

Thank you!

Erlend Ness Rannem
Trondheim, May 2021

Contents

List of Tables

List of Figures

1 Introduction

No language engages in direct mappings between thoughts and words. When producing a sentence, it is inadequate to select the relevant words of information and string them together in an order that conveys the meaning-relationship among them (Pinker, 1989). The sentence will turn out nonsensical, or partly comprehensible at best. Instead, the overall meaning of the sentence is computed compositionally through an interplay between each involved linguistic participant. (After all, this is their collective goal: to convey information.) Each participant serves as some sort of source of information in this regard. But their contribution differs. The participant that provides the most information is the one that denotes the situation described, the verb word. According to the widely accepted view of traditional generative theory (e.g., Chomsky, 1965), the lexical entry of a verb entails semantic and syntactic information. From which, a broad range of effects radiates out that are placed onto the verb's surrounding environment. This makes the verb choosy: choosy as to what sentences it appears in; choosy as to what participants it wants nearby; choosy as to what position the participants must occupy; and choosy as to what roles the participants must play out. Together, the verb and its arguments carry out what is called *argument structure*: the lexical representation of a verb's arguments as specified by the properties of the verb.

As such, the verb is the nucleus of the sentence and represents the lexical unit around which everything revolves. The properties it bears determine the argument structure it brings about, which in turn acts as the driving force of constructions (Koenig et al., 2003). However, a verb's properties of semantic and syntactic information are not of equal status in this regard. The syntactic expression of a verb's argument structure is outlined by the information about participants included in the lexical entry of this verb. Simplified: *meaning governs form*. Under this view, the outcome of what syntactic category and what semantic role a participant will be lexicalized as is therefore assumed to be pre-determined by the semantic properties of the verb. This gives rise to a predictive relationship between a verb and its arguments but also that of *sensitivity to argument structure*.

This sensitivity is rooted in the idea of the meaning expressed by a verb, together with its arguments, appears immediately accessible to linguistic intuition. Given the presumption that semantic properties of a verb pilot sentence structure, when a verb is identified during input processing, its lexical entry is activated upon to predict the upcoming post-predicate argument(s). Several studies have demonstrated sensitivity to argument structure by native speakers, wherefrom it is apparent that speakers exploit the lexically encoded information of a verb to make predictions on what entity is to follow (Altmann & Kamide, 1999; Kamide et al., 2003). This sensitivity seems to go as follows: a speaker identifies the constraints set by a verb's semantic properties (its *selectional requirements*) on the internal argument slot. She then makes predictions on what noun phrase (NP) would most likely follow given the semantic context. If her linguistic intuition is correct, she would show sensitivity in this regard.[1]

A question then arises whether this sensitivity to argument structure is exclusive to native speakers or if it is purely based on target language proficiency. Though scarce research has been devoted to this topic, the answer seems to be the latter. The findings from the collaborate project by Johnsen (2016) and Reine (2016), and the study by Hammerås (2017) indicate a direct relationship between proficiency and sensitivity. In all these studies, *verb constraint* (of different degrees) was the condition of interest: participants with sufficient proficiency would be able to make use of the semantically restrictive context posed by a verb to alter their behavior similar to that of native speakers. As such, verb constraint has been shown to be a key variable in predicting upcoming verbs. But does the variable of *verb frequency* influence as well?

The present project looks at to what extent Norwegian second language (L2) learners of English display sensitivity to argument structure, and how verb constraint and verb frequency influence in this regard. Three levels of verb constraint together with two levels of verb frequency constitute the six verb categories used in a sentence completion test. Participant responses were compared against native speaker responses. In addition, participants underwent a vocabulary test to establish the level of proficiency of the individual participant and the participating group as a whole.

The expectation was that vocabulary test scores would correlate with level of sensitivity. If a participant scored high in the vocabulary test, she would also perform more like a native speaker in the sentence completion test. This would show increased sensitivity to argument structure. If, on the other hand, she scored low in the vocabulary test, she was more likely to perform worse in the sentence completion test (less target response overlaps). She would be more susceptible to show signs of first language (L1) transfer; less likely to "detect" the semantic information stored in the verbs, and less likely to know the verbs' meaning(s). All things considered, she would show less sensitivity. Another expectation was that verbs with the

[1] Given that she, more often than not, successfully predict the upcoming input over a large set of test trials. The higher rate of success indicates a higher display of sensitivity.

most constraint would the most target responses due to the additional context provided by these (hence, reducing the number of possible NP-fillers). Following this train of thought, as constraint decreases and context diminishes, so would the number of target response overlaps as well. Also, verbs more familiar to the learner were expected to yield more target responses compared to less familiar ones. Therefore, frequent verbs were anticipated to perform the best due to their increased exposure (thus, enhanced familiarity) than infrequent verbs. Concurrently, verbs combining a constraint and a frequent nature were assumed to be the supreme verb group altogether.

2 Theoretical framework

The following chapter will present the theories and research most relevant to the thesis. The first sections will introduce the concept of argument structure and assert the appeal for a lexical-semantic stance (opposed to a syntactic one) relating to its realization. Afterward, what it means to be sensitive to argument structure, and how it has been demonstrated by L1 users will be presented. Finally, the thought-to-be problematic nature of L2 sensitivity with all its intricacies will be subject to discussion.

2.1 Before syntax, there is semantics

The present study is interested in word (or more specifically, verb) meaning and has for that reason based its theoretical framework within the field of *lexical-semantics*; i.e., the study of word meaning and meaning relations between words (Saeed, 2016). With the emphasis on verbs, they represent the linguistic unit that lexicalizes properties of happenings in the world and are *event descriptions* (Levin & Rappaport Hovav, 2005). As a verb names participants involved in its denoted event, this makes it the organizational core of any sentence. Taken one step further, it makes the meaning held by a verb essential to the overall sentence meaning. Following these ideas, the study at hand assumes that a verb's meaning encodes a specific conceptualization of an event and, in the words of Levin (1993), that "[…] the behavior of a verb, particularly with respect to the expression and interpretation of its arguments, is to a large extent determined by its meaning" (p. 1).

The hypothesis of syntactic properties of a verb being determined by its meaning has a long history in linguistics. The most outstanding illustration of the role of meaning regarding its deterministic character of syntax is the tendency for arguments bearing certain semantic roles to be associated with particular syntactic environments. This linkage of meaning and form has been noted by traditional grammarians dating as far back as Panini.[2] In modern linguistics, however, the unique role of the verb, in the context of realizing syntactic expressions, has been acknowledged by traditional generative theory – e.g., *Government and Binding* (Chomsky, 1981). According to such approaches, syntactic structures are generated from the level of words. That is, from the semantic aspect of a lexical unit. The idea is that lexical constituents are projected from a lexical core which entails that the lexical-semantic facet of lexical units is

[2] Indian grammarian who wrote the Sanskrit treatise on grammar in the 6th to 5th century BCE.

projected into the syntax. This gives rise to a semantic-driven grammar: all syntactic representations are derived from the *lexicon* (one's mental vocabulary). Under this view, every lexical unit is listed in the lexicon with information of its form and meaning. Verbs, as the organizational core of any sentence, express their participants' semantic roles and syntactic locations accordingly. This results in a verb and its arguments interconnecting, semantically and syntactically, in what is known as argument structure.

2.2 Argument structure and its two faces

Given that the verb word carries the information about form and meaning, it follows that its 'behavior' determines the structural organization of its sentence. According to classic linguistic theory (such as Chomsky, 1965; Katz & Fodor, 1963), there are separate stores of words (the lexicon) and word knowledge in the mind of a speaker, and that only certain semantic properties function within the former. A construct of this theory is the concept of selectional requirements (also called *selectional restrictions*) which refers to lexicon-internal constraints that verbs place on their arguments. These requirements not only determine what entities a verb must have in its presence (relating to its pre- and post-predicate positions), but the meaning of the verb determines what kind of arguments these entities must be realized as (Haegeman & Guéron, 1999). In other words, it is said that the verb has argument structure: the lexical representation of a verb's arguments specified by information such as the number of arguments, their syntactic expression, and their semantic relation to the verb which ultimately determine the syntactic environment of arguments. At this level of representation, verbs and the nominals derived from them share a common representation in the form of lexical entries distinct from that of syntactic categories (Butt, 2008). Consider the verb *kiss* in (1).

(1) The boy kissed the girl.

At the conceptual level, the verb *kiss* expresses an event that involves two participants: a '*kisser'* (*the boy*) and a '*kissee'* (*the girl*). These participants play both a semantic role and a syntactic role in the sentence. Semantically, they are classified according to *thematic roles*. These roles are specified in the verb's lexical entry: the verb *kiss* is a transitive verb that assigns two thematic roles, an AGENT (the 'doer' of the action) and a THEME (the 'receiver' of the action). (More on thematic roles in Section 2.3.) At the same time, the lexical entry of *kiss* also specifies how the participants will be realized syntactically. It requires a subject NP (the external argument) and a direct object NP (the internal argument). Note that the verb *kiss* ends up with two sets of categorizing labels of its participants, as illustrated in (2): one rooted in meaning (2b), another in form (2c).

(2) *Kiss*

a. '*Kisser'* and '*kissee'*

b. AGENT and THEME

c. Subject and object

The takeaway from (2) is that "[a]rgument structure has two faces, semantic and syntactic" (Bresnan, 1995, p. 1). As a result, it has been recognized that semantic and syntactic information, in relation to argument structure, must act separately (Tenny, 1994).[3] No matter where the line is drawn between semantics and syntax regarding their representation in argument structure, these will interconnect but not overlap. Thus, they must be treated individually. Still, no consensus has been reached as to what quality of the verb (semantic or syntactic information) that controls argument structure realization.

One approach which explores how meaning is derived from form (the syntactic frame) is the *exoskeletal model* by Borer (2005). In this context, lexical items contribute only for the sake of grammatically encyclopedic meaning. Instead, the structure is what determines the syntactic properties of all other aspects of meaning. This *neo-constructionist approach*[4] sees argument structure to be read directly off the syntactic structure, while the semantics is reduced to that of contributing to the content only. The pervasiveness of this language model is found in verbs with multiple meanings having multiple argument realizations.

(3) *Siren*

a. The factory horns sirened throughout the raid.

b. The factory horns sirened midday and everyone broke for lunch.

c. The police car sirened the Porche to a stop.

d. The police car sirened up to the accident site.

---

[3] The categorizing labels of thematic roles and grammatical relations occupy two separate domains of linguistics. Accordingly, there is not a one-to-one correspondence between the two. Although there is a tendency for the linguistic unit identified as the AGENT to occupy the subject position, or for the THEME to be located in the direct object position, there are examples in the literature that refute this. For example, Saeed (2016) points to the possibility of omitting the AGENT and have an INSTRUMENT occupy the subject position instead.

[4] "[A]n approach which shifts the computational load away from the lexical entry to the syntactic structure, subscribing to the view that an independent lexicon includes a minimal amount of structural information, and that it is structural constraints which determine traditionally lexical properties such as syntactic category type and argument structure […]" (Åfarli, 2007, p. 3).

e. The police car sirened the daylight out of me.

(Borer, 2005, p. 3)

Borer (2005) challenges the traditional approach of each verb having a structured lexical entry which alone determines the realization of arguments. Instead, she claims it is the syntactic expression of the arguments that determines meaning. In this sense, the approach uses elaborated syntactic structures accompanied by a reduction of the lexical item's attribution to a minimum. The lexical entry of a verb registers only its core meaning. It is only a verb's root (without any associated arguments) which is combined with the event-based meanings represented by the syntactic structures. This account on argument structure poses a valid concept on how form governs meaning. Still, flaws of this approach have been pointed out.[5]

A second approach, and the one the present thesis adheres to, is the lexical-semantic school. It holds that it is the meaning of a verb that accounts for how a verb's syntactic structure is played out. Information regarding the syntactic realization of arguments is situated in the verb itself by specifying the number of arguments the verb takes, their semantic nature, and the syntactic structure these semantic roles must hold. This assumption guided Levin (1993) in her systematic description of verb classes and argument structure alternations. In her listings of verbs, she shows the correlations between the semantics of verbs and their syntactic behavior by pointing to some facets of the semantics that govern syntax and the interpretation of arguments. From which, Levin (1993) argues that "argument structures might in turn be derivable to a large extent from the meaning of words" (p. 12). Moreover, she points to how native speakers can make correct judgments on what alternations verbs allow. Although two verbs might be closely related, speakers are aware they may not be adequate across alternations.[6] For example, speakers know that the verbs *fill* and *cover* allow one possibility of the locative alternations, whereas *dump* and *pour* allow the other.

(4) *Fill*

a. *Gina filled lemonade into the pitcher.

b. Gina filled the pitcher with lemonade.

(5) *Cover*

a. *Monica covered a blanket over the baby.

[5] See Levin and Rappaport Hovav (2005, p. 193) and Kiss (2011) for discussions.

[6] Levin (1993) defined about 200 verb semantic classes, where, in each of these, verbs share a certain number of *alternations* (e.g., locative, transitive, causative, etc.).

b. Monia covered the baby with a blanket.

(6) *Dump*

a. The farmer dumped apples into a cart.

b. *The farmer dumped the cart with apples.

(7) *Pour*

a. Carla poured lemonade into the pitcher.

b. *Carla poured the pitcher with lemonade.

(Levin, 1993, p. 2)

On the other hand, verbs like *spray* and *load* allow both possibilities. Based on such illustrations, Levin (1993) argues that it is essentially the meaning of the verb that enables speakers to make such subtle judgments. In this respect, verbs of similar meaning can be classified according to their shared semantic components and the type of event they denote. Verbs of the same semantically defined class will therefore express identical syntactic environments.

Later, Dimitriova-Vulchanova and Dekova (2007) provided another account to the "meaning governs form"-discussion. Here a verb's meaning is not the sole determining factor in how the syntactic landscape unfolds; rather, it is also contingent on the type of event lexically encoded in the verb. Following the representational format and assumptions of *The Sign Model* (Hellan & Dimitriova-Vulchanova, 2000) allow for a verb's denoted event to be represented on several distinct dimensions. This multi-dimensional system provides specifications of participants involved in the event. Each participant is assigned values on these dimensions that reflect various aspects of its involvement. One such dimension is that of *Force* which represents an event where the emission of physical force occurs. (The verb *kick* represents such an emission of force and will be used as an example in this regard.) Based on its role in such an event, a participant may be assigned the values of *Source* if the participant releases force (if the participant is the one who performs the kick); *Source Extension* if a part of the participant performs the action (the leg of the participants); and *Absorber* or *Limit* if the participant is the entity affected by the force (the object kicked). Other dimensions include *the Control-*, *the Monodevelopment-*, and *the Conditioned dimension*; each represents a specific type of event, and each event has its own set of values (for a full account on the dimensions, see Dimitriova-Vulchanova & Dekova, 2007; Hellan & Dimitriova-Vulchanova, 2000). Furthermore, several of the dimensions can coincide in verbs. This suggests that it is the chosen representational format that decides the participants' involvement in an event. As such, the syntactic realization of a verb and its arguments is determined by the meaning of the verb and the chosen frame of lexicalization.

2.3 Thematic roles

Several influential approaches to understanding argument structure have attempted to characterize semantic selection through the different types of roles each participant plays in an event. Lists of the types of roles arguments can manifest have varied from author to author, but the term *thematic roles* (Dowty, 1991) will be adopted for subsequent discussion. Drawing on the vast amount of work devoted to understanding the nature of semantic roles, it is possible to formulate a standard list of thematic roles as (8).

(8) Thematic role list

a. AGENT: the initiator of some action.

b. PATIENT: the entity undergoing the effect of some action.

c. THEME: the entity which is moved by an action.

d. EXPERIENCER: the entity which is aware of the action or state but which is not in control of it.

e. BENEFICIARY: the entity for whose benefit the action was performed.

f. INSTRUMENT: the means by which an action is performed or something comes about.

g. LOCATION: the place in which something is situated or takes place.

h. GOAL: the entity toward which something moves.

i. SOURCE: the entity from which something moves.

j. STIMULUS: the entity causing an effect in the EXPERIENCER.

(Saeed, 2016, pp. 150-151)

A widespread way of illustrating a verb and its thematic roles has been by a *thematic grid* (Williams, 1981). As a verb has particular (selectional) requirements for its arguments, the thematic roles associated with a verb are then listed with respect to the transitivity property of the verb and to what types of thematic roles its arguments must hold. By way of illustration, the verb *put* has the theta-grid of (9a), exemplified in (9b).

(9) *Put*

a. Put V: <AGENT, THEME, LOCATION>

b. [John]$_{\underline{\text{AGENT}}}$ put [his wallet]$_{\text{THEME}}$ [on the table]$_{\text{LOCATION}}$

Example (9a) shows that the verb *put* expresses an event that involves three participants, their thematic roles, and the positions of these in relation to each other and the verb itself. As exemplified in (9b), it has a putter, realized as an AGENT argument (*John*); the object being put somewhere, realized as a THEME argument (*the wallet*); and the place at which the object is put, realized as a LOCATION argument (*the table*).

The concept of thematic grids specifying the thematic properties of verbs and thematic roles being "determined completely and solely by verb meaning" (Dowty, 1991, p. 76) supports the hypothesis of verbs governing sentence structures. Subscribing to this approach involves the idea of lexically encoded information in verbs comprising the description of an event, the number of entities linked to this particular event, and the roles these entities play in this event. Building on these assumptions, theories of generative grammar estimate that the syntactic expression of arguments is also to be found in the lexical entries of verbs: i.e., in semantically based representations of argument-taking properties where the syntactic realization of these arguments is derived from these argument structures (Levin & Rappaport Hovav, 2005).

Thematic roles have not been an uncontroversial topic, however. This is reflected in the numerous attempts made to formulate sets of thematic roles that 'actually' work, and in the endeavor to integrate the lexical-semantic specifications with syntactic structure (Stringer, 2010). This led to Chomsky (1981) introducing the *Theta-Criterion*: "Each argument bears one and only one θ-role, and each θ-role is assigned to one and only one argument" (p. 36). Later, Baker (1988) introduced his mapping principle in this *Uniformity of Theta Assignment Hypothesis* (UTAH), claiming that NPs bearing identical thematic roles to a verb must be realized in the same syntactic relation to that verb. Nonetheless, despite many efforts attempting to make the concept of semantic roles work, there have been several examples in the literature that points to inadequacies of this approach. One of which being the lack of evidence for an underlying one-to-one correspondence between thematic roles and NPs. Jackendoff (1972), for example, points to sentences in which arguments seem to fulfill more than one role.

(10) John deliberately rolled down the hill.

He argues that (10) displays dual thematic role assignment due to the subject argument receiving the roles of AGENT and THEME: the NP *John* is simultaneously the entity initiating the action but also the one moved by it. On the other hand, and equally problematic, are cases in which verbs appear to hold two arguments bearing the same role. One of the examples illustrated by Dowty (1991) is that the external and internal argument of the verb *resemble* do not seem to be distinguished from one another due to there being "[…] no motivation for assigning distinct roles to them on semantic grounds" (p. 556).[7]

Thus the debate on the semantic definition of thematic roles remains still unresolved, and most linguists agree that the concept of thematic roles is problematic (Butt, 2008). One attempt aimed at solving the problems that has been widely attractive is a theory of Dowty (1991) whose object was to untwine the troublesome semantic role outlining by using a prototype conception. His revised representation of semantic roles focuses on a key problem for argument structure: *given a transitive verb, what determines which argument is its subject and which its object?* From the semantic point of view, a verb imposes entailments on its arguments based on the role each argument plays in the event described. Dowty (1991) sees these entailments as *lexical entailments* in that they derive from the verb alone – this means that context does not influence. In this regard, he is interested in the lexical entailments that make up the properties of subjecthood and objecthood. This leads to two blocks of property clusters.

(11) Proto-AGENT entailments
- a. Volitional involvement in the event or state.
- b. Sentience.
- c. Causing an event or change of state in another participant.
- d. Movement.

(12) Proto-PATIENT entailments
- a. Undergoes a change of state.
- b. Incremental theme.
- c. Causally affected by another participant.
- d. Stationary.

(Dowty, 1991, p. 572)

The idea is that every verb entails these properties which are checked against its arguments to see how well they correspond to the traditional roles of AGENT or PATIENT: the greater the number, the more typical an AGENT or PATIENT it is (Saeed, 2016). Thus, the innate flexibility in the proto-roles allows arguments to display a 'degree of membership' to either cluster. In terms of which argument that corresponds to what grammatical relation, Dowty (1991) introduces the *Argument Selection Principle* which states that it is the argument with the

[7] Dowty (1991) calls such verbs of such character *symmetric stative predicates*. Their thematic role assignment goes as follow: if this resembles that, then that resembles this (and vice versa). This means that with no apparent asymmetry between the two arguments, nothing drives role type assignment.

greatest number of Proto-AGENT properties that will be lexicalized as the subject; while the argument with the most Proto-PATIENT properties will be lexicalized as the direct object.[8] In this regard, the work by Dowty (1991) allows, through his take on semantic roles, predictions of the syntactic realization of a verb's arguments to originate in the information encoded in the verb itself. Regardless, for the sake of simplicity, the thesis will keep to the traditional view of thematic roles as presented in (8).

2.4 Predictability, context, and sentence processing

To understand the appeal for a lexical-semantically driven argument structure, the connection between predictability, context, and sentence processing needs to be clarified. According to lexical-semantic approaches to argument structure, a verb's semantic properties can be exploited to predict its syntactic argument structure realization (Gropen et al., 1991; Levin & Rappaport Hovav, 2005). It is however another layer to consider as context also affects meaning in that it contributes to the linguistic encoding of a situation. A verb may inherently hold multiple meanings, though context determines which meaning it should carry. The importance of context to determine the meaning of the event denoted by a verb is illustrated by the verb *run* in (13).

(13) *Run*

a. Sam runs.

b. Sam's nose runs.

c. Sam runs for president.

d. Sam runs his shoes to shreds.

The meaning of a verb description must be understood in relation to the context in which it appears. The verb *run* can be found in contexts describing the activity of moving one's legs rapidly (13a); the state of having excess nasal drainage (13b); one's attempt to become elected to a certain public position (13c); or the cause of change resulting from an activity (13d). At first glance, it seems that a verb's meaning is unrestricted, though a closer look reveals that this is not so. Several linguistic properties are governing the range of meanings associated with a verb which encompass entire semantic classes (Levin & Rappaport Hovav, 1998). As for the verb *run,* it is part of the verb class of manner of motion. Therefore, the range of meanings

[8] In addition to the argument selection principle, Dowty (1991, p. 576) presented a couple of ancillary principles (*Corollary 1 & 2*) and the characteristics of *Nondiscreteness*.

available to a verb seems to be determined by its semantic membership. In this regard, the role of context serves to disambiguate a verb's meanings down to a suitable one through the linguistic information co-occurring with the event described by the verb. According to this idea, context includes "[…] relevant knowledge available in the same place and time as an event" (Dimitriova-Vulchanova & Weisgerber, 2007, p. 55).

In relation to predictability, contextual information is believed to influence the processing of subsequent words (Van Petten & Luka, 2012). This view supports "a language processing system in which semantic interpretation, as well as syntactic processing, is conducted incrementally, with early integration of contextual information" (Sedivy et al., 1999, p. 109). Moreover, since language users interpret input continuously and incrementally (Kamide et al., 2003) they make use of all sources of information available during the processing of sentences (Matsuki et al., 2011). The most prominent of such information is embedded in the verb and the context it is presented within, to anticipate what lexical items will come next in a sentence. Similar to how a tennis player is constantly trying to predict where the opposing player will place the ball does a speaker somewhat the same in language processing. To follow through with the metaphor, both exemplars base their prediction on the available information: the tennis player on the position and playstyle of the opposite player; the speaker on his lexical knowledge of preceding words and the context of the event.

Although predictability is a controversial topic of debate[9], the behavioral measures of eye-tracking and event-related potentials (ERPs) have demonstrated that language users engage in predicting upcoming input. For example, studies have demonstrated that such prediction is found at different levels of language: semantics (Altmann & Kamide, 1999), syntax (Van Berkum et al., 2005), and form (Delong et al., 2005).

2.5 Sensitivity

2.5.1 L1 sensitivity to argument structure

The experiment that first established evidence for a predictive relationship between a verb and its arguments was that of Altmann and Kamide (1999). They found that the meaning of verbs can be used to predict thematic roles that are yet to be syntactically realized. Participants' eye

[9] "One might well predict that what comes after 'The big star's beside a little …' is likely to be a noun (though it might be BLUE or VERY OLD), but that still leaves open some tens of thousands of choices" (Jackendoff, 2002, p. 59).

movements were recorded while being presented semi-realistic visual scenes.[10] In one of which, *a boy*, *a cake*, and several distractor objects were visualized while participants listened to sentence (14) or (15).

(14) The boy will eat the cake.
(15) The boy will move the cake.

The verbs *eat* and *move* differ in terms of their degree of constraint (set by their selectional requirements). Whereas the former requires a THEME argument with the quality of being edible to occupy the ensuing syntactic slot, the latter needs a THEME NP that can undergo motion in this position. Comparatively, *eat* is the more constraining verb of the two in that its requirement is more specific. This was also what the gazes of the participants demonstrated: when presented with (14), participants would direct their gaze towards the appropriate object (*the cake*) before the target object had been uttered than when presented with (15) (Altmann & Kamide, 1999). The result seemed to be rooted in there being only one object in the scene that had the property of being edible, while several objects could undergo movement (e.g., *a toy car* and *a balloon*). In other words, only one object suited the selectional requirements of *eat* whereas several met those of *move*. The study showed that verbs that exhibit highly semantically constrained contexts allow participants to make predictions as to what post-predicate linguistic units will follow. Verbs with stricter sets of selectional requirements perform better in this regard.

Following this idea of semantic- and syntactic information being intertwined in a verb and its arguments, Friederici and Frisch (2000) conducted an ERP study to investigate verb- and argument-specific information in sentence processing. This was achieved by recording participants' brain activity while presenting argument structure violations. The violations were due to either a mismatch between either (16) the selectional requirements of the verb and the semantic features of the object NP, or (17) the verb and its number of arguments, or (18) the type of argument.

(16) *Anna weiß, dass der Kommissar (NOM) den Banker (ACC) abbeizte (V) und wegging.
*Anna knows that the inspector (NOM) the banker (ACC) stained (V) and left.
(17) *Anna weiß, dass der Kommissar (NOM) den Banker (ACC) abreiste (V) und wegging.
*Anna knows that the inspector (NOM) the banker (ACC) departed (V) and left.

[10] Multiple scenes were presented to the participants, accompanied by a sentence pair. In each case, the selectional requirements of the first verb in each sentence pair did only fit to one object, whilst the requirements of the second verb fit more than one object.

(18) *Anna weiß, dass der Kommissar (NOM) den Banker (ACC) beistand (V) und wegging.
*Anna knows that the inspector (NOM) the banker (ACC) helped (V) and left.
(Friederici & Frisch, 2000, p. 481)

The results favor the idea of language users incrementally processing input by inspecting each unit in an item-by-item manner. Elements involved in an argument structure must then satisfy structural and thematic requirements. The parser uses verb-specific information against which incoming arguments are checked, and argument-specific information against which the incoming verb is checked (Friederici & Frisch, 2000). Collectively, these allow for lexical integration. By building on an earlier idea by Osterhout et al. (1994) that the verb controls the syntactic realization on forthcoming arguments, Friederici and Frisch (2000) claim that any lexical item of an argument structure can be used to predict incoming input as both semantic- and syntactic information are encoded in each involved unit's entry.

Additionally, in response to the different violation conditions different ERP components were evoked. An N400 for a semantic violation in cases with an incongruency between the verb's selectional restrictions and its argument's semantic features; a biphasic N400-P600 for transitivity violations; and a P600 for a violation of the type of argument in cases of incorrect case marking (Friederici & Frisch, 2000). A similar isolated N400 effect was found by Frisch et al. (2004) when investigating the relationship between word category information and argument structure information of verbs during sentence parsing. They believe this N400 response can reflect the fact that a semantic violation impedes thematic role assignments, which further thwarts the derivation of the proposition of a sentence. Moreover, the study supports the idea that the lexical entry of a verb governs syntax as "the verb is typically structurally integrated […] before the parser identifies the verb's arguments and determines what the thematic interpretation is that they have to be provided with" (Frisch et al., 2004, p. 214).

Another ERP study on argument structure violations was conducted by Wassenaar and Hagoort (2007). They investigated online thematic role assignment during sentence-picture matching across three groups (of which the data from the group of healthy individuals is reported hereunder). Participants were first presented with a picture of an event (*a woman pushing a man in a wheelchair*) followed by an auditory sentence which either matched (*the woman pushes the man*) or mismatched (*the man pushes the woman*) the preceding scene. In the case of the mismatch, after hearing the sentence fragment "The man pushes", the NP *the man* could be realized as both the grammatical function of the subject and the thematic role of AGENT. Given the presence of a man and a woman in the scene, this was immediately followed by an automatic process of THEME role assignment mapped onto *the woman* – before the lexical item was uttered. The mismatch between the event structure of the picture and the auditory sentence was recognized upon processing the argument structure of the verb *push*, together with one of the thematic roles. This taxing process was visible in participants' brain waves; shown by an N400-P600 pattern.[11] When investigating the picture, the action describing an AGENT and a THEME forms a mental representation of the depicted event. Thus, while hearing the mismatching sentence afterward, a linguistic representation is incrementally formed as the words are processed. These representations are then compared in which a violation yields an N400. The fact that this negative peak effect was already seen as soon as the verb was heard suggests that thematic roles are assigned immediately as the argument structure is available. Not having to wait for the sentence to be completed implies that speakers use the context from the picture while processing the spoken sentence to detect discrepancies when confronting the verb (Wassenaar & Hagoort, 2007).

The predictive relationship between a verb and its arguments has allowed studies to exploit different aspects of argument structure to investigate reactions made by L1 speakers in response to different violation conditions. As the abovementioned studies have shown, L1 speakers react to subtle anomalies of the number of arguments, the type of arguments, and semantic violations. In turn, native speakers are sensitive to the argument structures of verbs. This sensitivity is rooted in the idea that a verb's argument structure is situated in its lexical entry. Thus, when the verb is identified during sentence processing, its lexical entry (with all its encoded information) is exploited to predict upcoming post-predicate arguments.

Although the preceding studies agree on sensitivity to argument structure by native speakers, the way in which this concept is acquired is another topic to consider. Pinker (1989) claims argument structure acquisition can be thought of in the following terms. During a child's first years, she is exposed to a finite number of sentences. However, since all languages are infinite, she must generalize from the exposures to an infinite set of sentences that both includes the input but also transcends it.[12] Consequently, this gives rise to an infinite number of hypotheses to how the target language works according to her input samples. All of these differ

[11] Wassenaar and Hagoort (2007) claim their study cannot explain the P600 effect, though they acknowledge the possibilities of 1) a mismatch between the participants' conceptual representation and the actual sentence; 2) a reassignment process of thematic roles; or 3) the strong thematic roles bias emanating from the picture, makes the grammatical role assignment process harder which the P600 effect reflects.

[12] Each language consists of a relatively small number of grammatical rules that allow the use of a finite set of words to create an infinite number of sentences.

from one another and from the correct hypothesis (the target language) in a way not detectable from the input samples alone. Considering the restrictions of the brain, the child must choose a set of hypotheses to follow where the learner can compare the hypotheses through the use of generalizations with the input so that the correct one can be extracted. In this sense, the child makes "[…] productive generalizations of many verbs to new argument structures, while excluding other verbs that are otherwise syntactically indistinguishable" (Pinker, 1989, p. 352) in her quest of acquiring argument structure. A successful acquisition is marked by the child having aligned herself with adults in that she has learned more and more accurate meanings for more and more verbs. Therefore, the occurrence of overgeneralization errors is thought to reflect the child's competence in that she has not acquired the correct meaning of the verbs yet. Still, research supports sensitivity to verbs at an early age (see Brooks & Tomasello, 1999).

2.5.2 L2 sensitivity to argument structure

Research affiliated to the area of L2 argument structure focuses broadly on how L2 learners come to acquire knowledge of the mappings between semantic representations of verbs and the syntactic configurations verbs will participate in, along with the accompanying arguments of verbs. The problem of L2 development of argument structure representations can be portrayed in the following terms:

> [I]f learners of English as a second language know that both *fall* and *drop* mean "to move downward", do they also know that "the apple fell to the ground", "the apple dropped to the ground", and "Sandy dropped the apple" are possible English sentences, but "*Sandy fell the apple" is not?
> (Juffs, 2000, p. 187)

This introduces the question of whether L2 learners can mentally represent the correct semantic features associated with individual verbs that may guide their mapping process between the meaning of verbs and their syntactic behavior.

Stringer (2010) argued that the actual problem of L2 acquisition of lexical semantics is situated in the L1 influence on the L2. He begins his take on the topic by acknowledging two insights gained by research into L2 lexical semantics. The first is that when undergoing interlingual translation, it is apparent that no exact one-to-one equivalence exists between languages in that either elements or word sense will be lost in translation. The second being that L2 learners draw from their L1 knowledge when making assumptions about novel L2 word meanings. Stringer

(2010) refers to the first of the challenges L2 learners face during argument structure sensitivity development as *lexical relativity*. Being built on the assumption of lexical equivalence being nonexistent, this concept holds that word meaning is relative to the meaning of the other lexical items that make up a lexicon. At a micro level, this means that the same speakers of a language may have different conceptualizations of states and events to a certain degree. As De Saussure (1959) points out, "[l]anguage exists in the form of a sum of impressions deposited in the brain of each member of the community […]" (p. 19). He is, however, careful to point out that native speakers do "produce – not exactly of course, but approximately – the same signs united with the same concepts" (De Saussure, 1959, p. 13). Although it is acknowledged that speakers of the same language may have different ideas of what makes up certain concepts, there is a greater gap between languages – i.e., at the macro level.

Languages do not lexicalize concepts the same way since there is no lexical equivalence cross-linguistically. A translation between languages may hold the same reference, but the languages almost always diverge in the senses related to the reference. For example, the English verb *drink* is used only for liquids, but one can "drink" smoke in Turkish; "drink" solid dose formulation in Japanese; and "drink" liquids and solids in Kazak (Stringer, 2010). Considering the mass of variation in the way argument structure is represented cross-linguistically (Bowerman & Brown, 2008), differing conceptualizations between languages may pose a problem for L2 learners if one follows the idea of verb meanings determining syntactic environments. Take argument structure as an example: since verb meanings denote happenings in the world, specify semantic roles, and the properties of participants, the verb selection will hinge on how a speaker conceptualizes particular events. Therefore, this act of selecting verbs to describe events is crucial as only certain verbs can describe specific events. Thus, L2 learners are susceptible to appointing an improper verb to express a happening if the verb is permitted for this representation in their L1 but not L2. Example (19) illustrates this phenomenon by using the Norwegian verb *ta* 'take' in a sentence possible in Norwegian but that requires another verb in English.

(19)

a. 'Sykepleieren	**tok**	en beslutning.'
b. *The nurse	**took**	a decision.
c. The nurse	**made**	a decision.

In this respect, lexical relativity is an organizing principle of the mental lexicon. This implies that speakers conceptualize events at a lexical level. It also points to such construals being language-specific to some degree. Therefore, L2 learners must tread carefully when making cross-linguistically syntactic generalizations on the basis of 'supposedly' equivalent lexical items.

The second of the insights mentioned above is termed *lexical transfer*. Based on the *Full Transfer/Full Access* model by Schwartz and Sprouse (1996), Stringer (2010) argues that L1 lexical semantics constitutes the initial state of L2 lexical acquisition. The assumption that implicit knowledge of the L1 is an important cognitive factor in the process of L2 learning can be regarded as an intuitive notion. Unlike infants learning their L1, L2 learning takes place among individuals who are already speakers of a language. When building their lexicon, L2 learners draw from their L1 knowledge of lexical items to map out the items' semantic and syntactic properties. Consequently, this gives rise to two problems for L2 learners. The first being that L2 learners, as pointed out by Stringer (2019), initially assume lexical equivalence due to the L2 lexicon build-up revolves around which L2 lexical items "correspond" to their L1 counterparts. Sprouse (2006) describes this as L2 learners maintaining the syntactic and semantic packaging of lexical items as listed in their L1 but are simply relabeling them with L2 phonology. This is a problem for L2 acquisition as a shift from one lexicon to another entails a legion of subtle differences in how we conceptualize the world and talk of it. The second being that the intuition to invoke information from the L1 in the L2 build-up gives L1 knowledge a deterministic role in how the L2 is unfolded. According to Schachter (1993), L2 learners face the task of learning a new language while equipped with a cognitive blueprint – comprised of their previous experiences as language speakers – of the possible shapes a language can take. This foreknowledge may prove misleading, on the other hand, making learnings forgo structural properties of the target language because it would violate rules in their L1. Going back to (19), Norwegian L1 speakers are prone to produce such English sentences if uncritical of the lexical transfer of *ta* 'take' into their English lexicon. Likewise, English learners of Norwegian would experience similar problems in reverse if uncritical of the verb *make* ("lage" in Norwegian) as illustrated in (20).

(20)

a. The nurse	**made**	a decision.
b. *'Sykepleieren	**lagde**	en beslutning.
c. 'Sykepleieren	**tok**	en beslutning.

Taken together, lexical relativity and lexical transfer constitute an apparent problem of cross-linguistic differences regarding alternations of argument structure. They can be depicted as two evils complementing each other: if the L1 lexicon serves as the initial state for L2 learners, the meaning of L1 verbs will be mapped onto the analogous L2 verbs, causing the L2 learner to permit the same behavior of the L2 verbs, even though the behavior might be confined to the L1 (Stringer, 2010). This prompts a vicious circle that is difficult to escape.

Thus far, problems connected to L2 acquisition of structures have been pointed out. Although there are certain recurrences in the mapping of argument structure to syntax, languages generally differ in their configurations of form and meaning mappings with argument structure as a focal point.[13,14] There is however a solution: extensive exposures to words across a range of semantic and syntactic contexts contribute to increased L2 proficiency (and therefore to the acquisition of L2 argument structures). The effect of frequency on L2 acquisition has caught the attention of researchers ever since Ellis (2002) raised the question about the relationship between frequency and second language acquisition (SLA): "How exactly does the frequency of patterns in the input affect acquisition?" (p. 165). This frequency refers to the relative frequency of linguistic features we as learners of a language are exposed to (Van Patten & Benati, 2010). The idea is that a language learner's knowledge of a linguistic construction depends on her experience of its use – a process thought to play a major role in SLA (see Larsen-Freeman, 1994; MacWhinney, 1999). As such, verb frequency is crucial. As frequency enhances a learner's knowledge of the properties underlying a verb's construction(s), then the greater reoccurrence of the verb in the input, the better. In this regard, there is a divide between high- and low frequency verbs (or *frequent verbs* and *infrequent verbs*, respectively). Based on such a divide, studies have shown that there exist frequency effects. Compared against infrequent verbs, frequent verbs are processed faster in picture studies (Oldfield & Wingfield, 1965) in word reading studies (Forster & Chambers, 1973), and in spoken word duration studies (Wright, 1979). In addition, frequent verbs are acquired faster (Treffers-Daller & Calude, 2015)

---

[13] Pinker (1989) claims that linking rules are "[…] near-universal in their essential aspects and therefore may not be learned at all" (p. 248). He offers an explicit proposal for the mappings from semantic roles to syntactic positions, as follows:

i. Link the agent to SUBJECT
ii. Link the patient to OBJECT
iii. Link the theme argument […] to SUBJECT unless SUBJECT is already linked; to OBJECT otherwise
iv. Link the goal to an OBLIQUE […] argument
v. Link the theme argument in a CAUSE TO HAVE predicate to the second object in a ditransitive construction.

(Pinker, 1989, p. 74)

[14] [T]here is an overwhelming tendency, cross-linguistically, for agents to appear as subjects and themes as direct objects, with other arguments appearing in oblique cases" (Naigles et al., 1993, pp. 136-137).

which suggests that their meaning (and therefore their "form-meaning mapping" pattern) are better known.

Building on the assumption of native-like structure competence being achievable, Johnsen (2016) points out that "[t]o what extent [L2 learners] are sensitive to argument structure in their language processing […] is unknown" (p. 11). Correspondingly, Johnsen (2016), with the help of Reine (2016), sought to map out to what degree Norwegian L2 learners of English display sensitivity toward argument structure. Their collaborative project used eye-tracking to measure the gaze behavior of participants toward appropriate target objects depending on the verbs' level of (semantic) constraint. To paint a picture of the L2 sensitivity landscape, the study focused on two groups that differed in their level of L2 proficiency. The findings clearly show that L2 learners display sensitivity to argument structure but that the magnitude of this sensitivity is contingent on proficiency. Similar results were replicated later by Hammerås (2017) who found a parallel between a participant's level of proficiency and her level of sensitivity: simultaneously as proficiency increases, so does one's sensitivity to the constraint of verbs which in turn accounts for one's sensitivity to argument structure.

3 Method

This experimental research project opted for a quantitative approach to test the attitudes of the target population in response to the specific research objective at hand:

> *the extent to which native speakers of Norwegian are sensitive to English argument structure, and how verb constraint and verb frequency influence in this regard.*

The experimental design consisted of an online survey comprised of a sentence completion test and a vocabulary test, wherefrom two sets of data analyses were initiated. One analysis concentrated on the collective performance on the incomplete sentences; another on the correlation between individual proficiency levels and performance to check for an interdependency connecting sensitivity and proficiency.

The following chapter will, in the first place, provide a closer description of the shaping of the material used in the incomplete sentences with regards to the properties of the verbs. Afterward, a discussion on the nature of the two online language tests is presented, followed by a description of the target population. Lastly, the procedure of the online survey will be provided.

3.1 Preparatory stage

Whereas the vocabulary test was standardized, the same was not the case for the sentence completion test which had to be built *ab initio*. The primary endeavor was to decide upon what verbs to include and why. The study settled on the two nominal variables of constraint and frequency from which verb categories were developed. Potential verb candidates to be included were required to manifest these features (though in different degrees).

3.1.1 Verb constraint

To reiterate, selectional requirements determine verb constraint. However, verbs rarely behave similarly – and the aspect of constraint is no exception to this norm. As previous studies have shown (e.g., Altmann & Kamide, 1999), some verbs impose several selectional requirements on their internal argument slot(s), whereas others are less strict in this regard. This assumption of different verbs displaying different degrees of constraint constitutes the underlying premise for the verb categories used in this study. That is, the constraint imposed by a verb can be exploited to predict upcoming input. A more constraining verb (one that places several selectional restrictions on the argument) would narrow down subjects' responses into a uniform – or close to uniform – response. This would be a consequence of the relatively few NPs compatible with the many required features. On the contrary side, a less constraining verb (one that sets few selectional restrictions) would expect to receive more unique responses since there are a plurality of potential argument fillers that may follow the verb. Furthermore, since a gradient view of constraint is assumed, it presupposes the existence of a middle ground level: a class of moderately constraining verbs that fits neither of the extremes. This three-way constraintness concept is transparent when comparing the argument-taking properties of (21)-(23).

(21) *Milk*

a. [PATIENT] argument.

b. Concrete entity.

c. Animate entity.

d. Mammal entity.

e. Non-human entity.

f. Produces milk.

(22) *Eat*

a. [PATIENT] argument.

b. Concrete entity.

c. Non-animate entity.

d. Characterizes as solid food.

(23) *Memorize*

a. [THEME] argument.

b. Abstract or concrete entity.

The highly constraining verb *milk* enforces multiple semantic restrictions on its following argument; all of which to ensure that the NP holding the argument slot is in fact 'milkable'. This implies that the entity must contain the semantic features of being an animate, non-human entity belonging to the mammal family (e.g., *cow*). On the other hand, a verb like *eat* is less discriminating as it is contempt with any concrete entity characterized as 'eatable'/edible (e.g., *food*). Finally, a verb such as *memorize* imposes little to no selectional requirements on the following argument seeing that any abstract concept or concrete entity can fill this argument slot (e.g., *number*) – indeed most things are 'memorizable'. These differing levels of the constraint of argument structure were used as the basis when categorizing verbs according to their restrictive nature. This led to the development of three categories of verb constraint: *the most-*, *the moderately-*, and *the least constrained category*.

However, listing a verb into its appropriate group based on its level of constraint proved to be difficult. The solution was to examine how restrictive a verb's argument structure is by looking at the number of NP-fillers it commonly takes. Whereas some verbs are strongly associated with certain argument fillers, other verbs are inherently more open as to what NP will follow. Put differently, the assumption was that the number of unique NPs a verb takes decreases parallelly as its level of constraint increases. Based on this, corpus linguistics was identified as the best approach. Three large-scale corpora were utilized: *the iWeb Corpus* (https://www.english-corpora.org/iweb/), *the Corpus of Contemporary American English* (https://www.english-corpora.org/coca/) (henceforth, COCA), and *the British National Corpus* (https://www.english-corpora.org/bnc/) (henceforth, BNC). By means of the fixed corpus syntactic construction *VERB the/a/an NOUN*, the searches took place in the following procedure. For any verb, two queries of corpus syntax were carried out. The first captured the independent frequencies for a verb and its most common NP-filler (e.g., *sow_v the/a/an NOUN*), whereas the second checked for the frequency of a verb and its most common NP-filler relative to the frequencies of competitor NP-fillers (e.g., *sow_v the/a/an seed*). Accordingly, a verb belonging to the most constrained category was presumed to have an NP that makes up the preponderance portion of the total number of NPs; a moderately constrained verb would have a declined portion of the total NP occurrences; and a least constraint verb would have an even minor portion.

The final step was to calculate how much of the total occurrences of NPs was made up of the most common NP-filler for the verb. To follow through with the *sow* example, the first search revealed that *sow* has 179 NP occurrences in the COCA corpus, whereas the second search showed that of these 170 occurrences 131 were *seed(s)*. A calculation[15] shows that the NP *seed(s)* makes up 73.18% of the total number of NP-fillers for *sow*. This result is consistent with the other corpora as well – 71.08% in the iWeb corpus and 85.71% in the BNC corpus. As a result, the data indicates that *sow* belongs to the most constrained category.

[15] (most common NP-filler * 100) / total number of NP-fillers

3.1.2 Verb frequency

Given that language learners are sensitive to the frequency of occurrence of linguistic units in the input (Ellis, 2002), the property of verb frequency was also expected to affect participants' responses. Rooted in the lexical-semantic idea of how information regarding argument structure realization is situated in verb entries (Koenig et al., 2003), this can only be exploited *if and only if* a learner knows the verb. But no language user – native or not – has knowledge of every verb in a language; rather, only verbs familiar to the speaker can be assumed to be known. In this regard, verb frequency and verb familiarity may be viewed as two sides of the same coin in that familiarity is measured by the frequency with which a verb occurs in a large-scale corpus as a representation of a language's general use.

Following this train of thought, a frequently occurring verb will by nature be familiar to the speaker. If familiar with a verb, the speaker is aware of what properties an NP to follow a verb must have to abide by the argument structure of the verb. Meanwhile, a low frequent verb is presumably less familiar due to its minimal appearance in the input. As a result, the properties the verb place on the following NP are to a lesser degree known which makes the learner more prone to insert an inappropriate filler. On that premise, a speaker is expected to perform better when continuing a sentence containing a frequent verb as opposed to an infrequent one. This frequency account can in turn explain the extent to which information is stored in the mind of the speaker. As Langacker (1987) lays it, frequency relates to the concept of entrenchment in cognitive linguistics in that the more frequent a verb is, the stronger it is *entrenched* and conventionally well established. From this point of view, high frequency of use reinforces semantic and contextual knowledge of the verb – thereby exceeding that of low frequency verbs. With this in mind, the expectation was that frequent verbs would generate few unique responses due to them being more known and thus encourage uniform answers; whereas low occurring verbs would promote an increase in unique responses due to their lesser recognized nature.

In classifying verbs based on frequency, the corpora were again visited. Not to be confused with the frequency of an NP following a verb (as discussed in the preceding section), this frequency refers to the number of times a verb occurs in a corpus. However, simply finding a verb's raw frequency in the corpora was not sufficient as the datasets are of vastly different sizes. Therefore, a common ground on which to compare the data was required. This was achieved by calculating *the normalized frequency*: the raw frequency of a verb in a corpus relative to the total number of tokens in that corpus.[16,17]

Once more, *sow* will illustrate the searches. The first objective was to find its raw frequency in each corpus. A basic search (sow_v) revealed that the verb *sow* occurs a total of 1 567 times in the COCA corpus, 31 609 times in the iWeb corpus, and 188 times in the BNC corpus. Having calculated how often *sow* occurs on average per million items shows that it has 1.56 instances per million (ipm) in the COCA corpus, 2.25 in the iWeb corpus, and 1.88 in the BNC corpus. The final step was to decide on a threshold regarding what ipm value should draw a wedge between a frequent verb and an infrequent one. Research on lexical frequency usually looks at general word frequency under the same scope; including all tokens of a language together into one list. Contrary to this, the present study looks at verb frequency in isolation and in a binary fashion where a verb is defined as either having high- or low frequency. As a result, the value that represents the wedge between the two sides was set to 10 ipm. Accordingly, *sow* was identified as an infrequent verb. The complete list of the verbs' raw and normalized frequencies is documented in Appendix 1.

3.1.3 The verb categories

A network of verb categories was established from which sentence fragments were produced by following the fixed sentence construction in (24)

(24) SUBJECT – VERB[+ past] – ARTICLE – [blank] – ADVERBIAL

This network is organized according to the variables of semantic constraint (*most*, *moderate*, and *least*) and verb frequency (*frequent* and *infrequent*). Interweaving these conditions in a binary manner produced a total of six categories, shown in Table 1.

[16] (number of hits in the corpus * base of normalization) / total number of items = ipm

[17] Due to all three corpora being large datasets, the base was set to frequency per million.

Table 1

Crosstabulation of verb categories

		Variable of constraint		
		Most constrained verbs	Moderately constrained verbs	Least constrained verbs
Variable of frequency	Frequent verbs	The most constrained and frequent category	The moderately constrained and frequent category	The least constrained and frequent category
	Infrequent verbs	The most constrained and infrequent category	The moderately constrained and infrequent category	The least constrained and infrequent category

The categorization resulted in 18 verbs in the most constrained and frequent category; 14 in the most constrained and the infrequent category; 20 in the moderately constrained and frequent category; 17 in the moderately constrained and infrequent category; 17 in the least constrained and frequent category; and 11 in the least constrained and infrequent category. From each category, 10 verbs were selected for use in the experiment – amounting to 60 verbs in total. The verbs that made up each category are listed in Appendix 2.

3.2 The sentence completion test

The sentence completion test has a long history in psychology. Ebbinghaus is credited with the development of the method in 1879 and used it to study the intellectual capacity and reasoning skills of children (Hersen, 2003). Over the years, the sentence completion test gained ground as a valuable method that authorized numerous iterations, e.g., *The Tendler Sentence Completion Test* (Tendler, 1930), *The Sentence Completion Test for the Office of Strategic Services Assessment Program* (Murray & MacKinnon, 1946), and *The Rotter Incomplete Sentence Blank* (Rotter, 1951).

More recently, Hamberger et al. (1996) adopted the method in the field of linguistics. They checked for how the degree of constraint imposed by a semantic context predicts the upcoming response. That is, a more distinct context should elicit few appropriate responses while an open-ended sentence will naturally allow several appropriate responses. Similarly, the purpose of the sentence completion test in the present project is to examine L2 sensitivity to the argument structure of verbs. Based on the predictability relationship between a verb and its following argument, sensitivity is elicited if participants display a native-like behavior by successfully matching the target response on the basis of the information stored in the lexical entry of the verb. When providing an NP-filler, the recognition of the verb will activate an item that is most typical to occupy the argument slot.

3.3 The lexical decision test

Due to the multifaceted nature of what it means to ‘know’ a word, the conundrum of how to measure L2 vocabulary knowledge is an ongoing debate (Enayat et al., 2018; Yue & Fan, 2016). Still, two dimensions of vocabulary knowledge that together seem to correlate with general language competence are that of *breadth* and *depth*. Vocabulary breadth relates to the size of word knowledge a learner knows (i.e., one’s mental lexicon); while vocabulary depth refers to the quality of word knowledge (i.e., how well words are stored in the mental lexicon) (Gyllstad, 2013). One particular type of YES/NO lexical decision test that seems to capture these aforementioned components adequately is the *Lexical Test for Advanced Learners of English* (henceforth, *LexTALE*: Lemhöfer & Broersma, 2012). Any experiment involving L2 users must control for participants’ proficiency level. However, instead of solely basing the mapping of competence on self-rating of proficiency and language background questionnaires, LexTALE is a practical and valid tool that allows for the measuring of English vocabulary knowledge. This estimate is in turn a good predictor of general English proficiency.

3.4 Participants

The target population was native speakers of Norwegian with English as their L2 who were attending some form of higher education at the time of participating. The online questionnaire was originally completed by a total of 65 participants; all of whom were recruited through social media. Of the original 65 participants, 11 (16.9 %) were excluded for various reasons: 1 (1.5%) participant was not a native speaker of Norwegian, 1 (1.5%) was not a university student, and 9 (13.9%) answered disingenuously. The final number of participants was 54 with 43 being males and 11 females (Table 2); the mean age was about 25 years (ranging from 19 to 38). All participants had Norwegian as their native language and English as their L2.

Table 2

Profile of participants (gender)

Gender	No. of participants	Percentage
Male	43	76.6%
Female	11	20.4%
TOTAL	54	100%

The study was reported and approved by *Norsk Senter for Forskningsdata* (NSD). No sensitive personal information was collected. Still, each participant was assigned a participant number, and the link between a participant's data and the corresponding number was only known to relevant research personnel. Finally, participation was voluntary, and no reward was offered.

3.5 General procedure

The experiment was carried out using *Nettskjema* (https://nettskjema.no/), a tool for designing and conducting online surveys. I opted for an online form of study due to its practicality as it enables researchers to test a larger number of participants than any "conventional" experimental setting would. A further reason was the motivation to continue practicing social distancing at the time of testing given the outbreak of Covid-19 at the time.

3.5.1 Part 1: Structure of the online study

Upon entering the survey, prior to any testing, (potential) participants were met with a welcome page that shortly described the study, its purpose, and the language tests. The subsequent page provided a thorough information sheet (see Appendix 3). Here details already given were expanded upon, in addition to more practical information – e.g., what an involvement in the study would mean, how personal information will be treated, and so on. Based on this outline of the research project, participants interesting in partaking signed a consent form (see Appendix 4) which granted the researcher permission to involve their data in the study. Individuals that signed could advance to the next page, whereas those that were not interested were instructed to close the webpage. This next section was devoted to painting a picture of the test takers through the means of questions. There were two sets of questions, each with its area of interest. One on personal information; another on participant's linguistic background (see Appendix 5).

One consideration of the introduction section was that it was written in Norwegian. This was done deliberately to ensure that participants understood what a commitment to the study implied. However, to prevent an immediate shift from L1 to L2 – when going from the introduction part to the language tests – an intermediary stage was added. This English written page simply informed that the testing would begin. This page intended to mark a switch from one language to another, and thus mentally prepare participants to switch the L2 lever on.

3.5.2 Part 2: Language tests

Then the actual testing took place. In the sentence completion test, participants were presented with several sentence fragments consisting of a subject, verb, and adverbial, but no object – e.g., *Kevin bought… last Monday*. The task was straightforward: make the sentence complete by inserting an appropriate object entity. For the responses to be in agreement with the corpora searches (in terms of syntax), participants were instructed to always start a response with an article followed by a single noun. Furthermore, to promote natural responses, participants were encouraged to not overthink their responses and avoid long phrases.

Afterward, the vocabulary test was presented. Following the procedure of Lemhöfer and Broersma (2012), the stimulus list was presented to the test-takers in a fixed manner. A series of strings of letters that were either words or nonwords were put forward, one at a time. For each stimulus, participants were asked whether this was an existing word or not. This was done by ticking off the accompanied YES-box if they believed it was a real word, or the NO-box if they did not. In case of doubt, participants were instructed to respond NO. However, if they were sure the word existed, even though they did not know its exact meaning, they could still respond YES. Important to note, since the results of this test would only be useful if responses were sincere, participants were encouraged not to look up the items in a dictionary.

4 Material

4.1 The sentence completion stimuli

Based on the 60 verbs deemed appropriate to serve as the research stimuli, 60 stimuli sentences were formulated (see Appendix 6). Every verb category was represented by ten sentences that reflected the inclusion criteria of the relevant category (regarding constraint and frequency). For instance, the most constrained and frequent category had only verbs with strict selectional requirements that occurred at a high frequency. The sentences ranged between 4 to 9 words long (mean: 5.68) and followed the fixed sentence construction shown in (24).

The three degrees of constraint, identified as the most-, moderately-, and least constrained categories, served different roles in the sentence completion test. The most- and moderately constrained sentences constituted the stimuli of interest. Verbs that belong to either category encode a narrow semantic nature (through selectional requirements) that requires a specific response to succeed – a response proficient L2 learners are prone to recognize. On the other hand, the broad/open nature of the verbs that make up the least constrained sentences may select a variety of potential NP-fillers. Few selectional requirements placed onto the argument slot entails that a mass of NPs is legitimate to be selected. Accordingly, no clear information of the verbs will help participants to specify a single adequate response. For this reason, the least constrained categories (regardless of frequency variable) performed the role of baseline verbs – verbs that were believed to be neutral with regards to argument selection.

In sum, 40 sentences formed the research stimuli sought to evoke reactions from participants that would provide insight on whether or not L2 learners are sensitive to argument structure. That is, will participants be capable of correctly responding with an appropriate NP-filler by the means of the lexical cues presented by the semantic constraint posed by a verb onto the post-predicate argument slot? In addition, 20 sentences functioned as baselines against which the other categories could be compared.

4.1.1 Selectional requirements

The following section provides the descriptions of the selectional requirements for each of the stimuli verb used in the sentence completion test. These definitions provide an abstract measure of what the verbs are most likely to take in relation to the properties of the argument. Note that only the verbs of interest (i.e., verbs belonging to the most- and moderately constrained category) are included below seeing that these display prominent selectional requirements – contrary to the least constrained verbs which express an open character with few to none demands.

[…] <u>*ate*</u> *the/a/an* [THEME argument] […]
A verb of ingesting which involves consuming solids. The verb will choose any entity that can be characterized as solid food (e.g., *cake*).

[…] <u>*boarded*</u> *the/a/an* [LOCATION argument] […]
A verb of motion referring to the act of entering a vehicle. Any public transportation vehicle is possible, though the named action is most commonly associated with the NP *plane*.

[…] <u>*boiled*</u> *the/a/an* [PATIENT argument] […]
A verb of preparing which shares properties with both change of states verbs, and verbs of creations and transformation. It expresses the preparation of food and is therefore most commonly followed by an edible entity; be it solid (e.g., *eggs*) or liquid (e.g., *water*). The verb can also be used as a causative construction in which case *kettle* is a likely NP contender.

[…] <u>*buttoned*</u> *the/a/an* [PATIENT argument] […]
A tape verb which describes the act of attaching one thing to another by passing a button through a buttonhole. Any clothing entity that can be closed with buttons can be selected (e.g., *a shirt*).

[…] <u>*cited*</u> *the/a/an* [PATIENT argument] […]
A verb of transfer of message that takes an NP that refers to the source of some (scholarly) work. The nature of the verb requires either the author of a work (i.e., a human entity) or the literature work of an author (e.g., *a book*).

[…] <u>*drank*</u> *the/a/an* [THEME argument] […]
A verb of ingesting which involves consuming liquids. This verb will select any drinkable liquid as its argument (e.g., *water*).

[…] <u>*decorated*</u> *the/a/an* [PATIENT/LOCATION argument] […]
A fill verb which describes the resulting state of a location as a consequence of putting something on or in it (Levin, 1993). The verb will most likely choose an NP for its argument slot that is either a location (e.g., *a room*) or any concrete entity associated with the named action (e.g., *a cake*).

[…] *dialed* *the/a/an* [PATIENT argument] […]

The verb describes the action of making a telephone call to a particular number. Thus, the most likely NP to fill the argument slot is indeed *number*.

[…] *dug* *the/a/an* [THEME argument] […]

A verb of creation and transformation that takes an affected object (i.e., an NP that refers to the created object). Importantly, the raw material (e.g., *gravel*) from which the final product (e.g., *a hole*) was created cannot be expressed as the direct object.

[…] *elected* *the/a/an* [BENEFICIARY argument] […]

An appoint verb that, when used in the syntactic frame NP VP NP, takes a direct object that refers to the official position or some other title of a person (e.g., *president*) (Levin, 1993).

[…] *entertained* *the/a/an* [EXPERIENCER argument] […]

An amuse verb that describes the change in the psychological or emotional state of the experiencer (Levin, 1993). This direct object entity is animate, and most likely a human or a collective group of humans.

[…] *extinguished* *the/a/an* [PATIENT argument] […]

A verb of destruction that relates to the total extermination of an entity. The direct object is most likely something concrete that can be ceased to burn or shine (e.g., *a fire*).

[…] *fired* *the/a/an* [PATIENT argument] […]

The verb may be a verb of throwing describing the instantaneous discharge of a projectile from a firearm, in which any weaponry that can set off a missile (e.g., *a cannon*) can fill the argument slot. Otherwise, it is a verb of removing in which it describes a person's dismissal from a working position, in which any human entity that is employed is expected (e.g., *an employee*).

[…] *flushed* *the/a/an* [PATIENT argument] […]

A wipe verb which relates to the removal of something from a surface by causing quantities of water to pass through it. The most common NP candidate is *toilet*.

[…] *fried* *the/a/an* [PATIENT argument] […]

A verb of preparing that describes the creation of a product (usually) through the transformation of raw material. More specifically, it involves the preparation of food that is best followed by the raw material. The named activity may select any NP that is solid, edible and makes up the raw material of the action (e.g., *chicken*).

[…] <u>*hammered*</u> *the/a/an* [PATINET/THEME argument] […]

A verb of contact by impact which describes the manner in which an entity is moved to bring it into contact with a separate entity of other characteristics (Levin, 1993). Seeing that the verb indicates the instrument used for carrying out the named action, it is most likely to take an NP that is associated with the instrument (e.g., *a nail*).

[…] <u>*killed*</u> *the/a/an* [PATIENT argument] […]

A murder verb that lexicalizes nothing about the purpose or manner of killing but describes the cause of death of a living entity. The verb may take any animate entity as its argument, though the most common argument filler is an all-encompassing noun that refers to a human entity (e.g., *man* or *woman*), or a typical animal to hunt when referring to an animal entity (e.g., *deer*).

[…] <u>*landed*</u> *the/a/an* [THEME argument] […]

The verb describes either the action of bringing an aircraft to the ground, in which case any type of vehicle that can fly (e.g., *a plane*) may fill the slot; otherwise, it is likely to be a verb of obtaining in which it denotes a person's success in achieving something desirable, in which case an abstract concept like *a job* or *promotion* is anticipated.

[…] <u>*lighted*</u> *the/a/an* [PATIENT/LOCATION argument] […]

A change of state verb that describes the creation of a product through the transformation of raw materials. The direct object can either refer to the raw material (e.g., *wood*) used to create the product, or the product itself (e.g., *a fire*) (see Levin, 1993).

[…] <u>*locked*</u> *the/a/an* [PATINET argument] […]

A tape verb describing the fastener/instrument used to attach one thing to another. Any entity with a locking mechanism may fill the argument slot, though *a door* is the most likely NP.

[…] <u>*married*</u> *the/a/an* [PATINET argument] […]

A verb of social interaction where the direct object participant must be of comparable status to the subject participant (i.e., *human*). Although the "marrying" action may occur homogeneously (i.e., same-sex marriage), the most likely direct object would be any label of the opposite gender of the subject participant. In this case, if the subject is *man*, the object is *woman* – and vice versa.

[…] <u>*milked*</u> *the/a/an* [PATINET argument] […]

A verb of possessional deprivation that is based on a noun considered to be an inalienably possessed part of an animal. As a result, the verb must be followed by an animal that contains this property of producing milk (e.g., *a cow*).

[…] _mowed_ *the/a/an* [PATIENT argument] […]
A verb of carving which includes the specification of the instrument used to carry out the action (i.e., a yard tool like string trimmer). This verb will most likely choose an NP that describes a ground covered with grass (e.g., *a lawn*).

[…] _ordered_ *the/a/an* [THEME argument] […]
A verb of obtaining that takes an argument of concrete entity that can be claimed ownership to through the process of transferring money. Most likely to select something that characterizes as edible (either a solid or liquid entity).

[…] _planted_ *the/a/an* [THEME argument] […]
Although the verb carries multiple senses, it is frequently used to lexicalize the act of setting something into the ground for the purpose of growth. Accordingly, any entity capable of growth is a strong contestant to fill the argument slot (e.g., *a tree*).

[…] _predicted_ *the/a/an* [STIMULUS argument] […]
A verb of future which expresses a person's estimate that a specified happening will occur in a future point of spacetime. The verb is most likely to take an event like *future* or *catastrophe*.

[…] _pruned_ *the/a/an* [PATIENT argument] […]
A carve verb which involves the notions of contact and effect on the entity undergoing the action. The argument slot will most likely be occupied by a member of the vegetable kingdom (e.g., *a bush*).

[…] _published_ *the/a/an* [PATIENT argument] […]
A create verb which describes the act of making something generally known to the public. Seeing that the verb takes an NP that refers to the created product as its direct object, the verb usually selects a literary work (e.g., *a book*).

[…] _pushed_ *the/a/an* [THEME argument] […]
A carry verb related to the movement of an object through accompanied force. Although many entities fit with the verb, it must be of such character that it is capable of moving as a response to the force administered by the agent. The most common NP to fill the argument slot is therefore a modest-to-small, concrete entity (e.g., *a button*).

[…] _read_ *the/a/an* [THEME argument] […]

A verb of learning that describes the acquisition of written information. When used in the transitive, the verb takes a direct object that qualifies as a typical object of the verb. Accordingly, this may be any form of literature or written text.

[…] _saddled_ *the/a/an* [PATIENT argument] […]

An equip verb which specifies what is provided in the event described. The verb is most likely to choose an animal entity commonly used for riding (e.g., *a horse*).

[…] _sailed_ *the/a/an* [THEME/LOCATION argument] […]

A manner of motion verb that describes a movement through the use of a particular type of vehicle. Any concrete entity associated with the named action of traveling across water can be chosen (e.g., *ship*). The verb might also describe the location at which the traveling action takes place. In which case, any term for an area of water is possible to fill the argument slot (e.g., *sea*).

[…] _smoked_ *the/a/an* [PATINET argument] […]

The verb usually refers to the act of inhaling and exhaling the smoke of tobacco. The most likely argument filler is an entity containing tobacco (e.g., *a cigarette*). The verb can also be seen as a verb of preparing in which food can be smoked to add flavor to the product. Although any entity characterized as solid food is possible, the most likely NP associated with the verb is *fish*.

[…] _solved_ *the/a/an* [PATIENT argument] […]

A verb that refers to the act of answering or explaining a problem. This verb is compatible with any abstract entity referring to a situation that is believed to be difficult in some way (e.g., *a mystery*).

[…] _sowed_ *the/a/an* [PATIENT/LOCATION argument] […]

Part of the spray/load verbs which, in this case, relates to the covering of land, earth, etc., for growth. As the verb will choose an entity that promotes growth, the most likely candidate is *seed*.

[…] _threw_ *the/a/an* [THEME argument] […]

A verb of throwing that describes a cause of motion event. The internal argument refers to the concrete entity that is set into motion by the agent. Thus, the direct object must be of a character (in terms of size and weight) that allows the agent to impose a force onto it which makes it move through the air. The most likely NP is *a ball*.

[…] *tied the/a/an* [PATIENT argument] […]
A tape verb that takes as a direct object an NP that names the fastener of the action. Will select an entity that can be fastened together or made into a knot.

[…] *washed the/a/an* [PATIENT/THEME argument] […]
This wipe verb involves the act of removing things from the surface of an object. Although many entities can be select as its argument, the most typical items are either a body part (e.g., *hands*), or any container for cooking/serving food (e.g., *dishes*).

[…] *watered the/a/an* [PATINET/LOCATION argument] […]
Belongs to the class of butter verbs whose meanings can be paraphrased as "put X on (something)" – in this case, to supply with water. The verb selects any entity belonging to the vegetable kingdom (e.g., *a plant*), or a location where something grows (e.g., *a garden*).

[…] *wrote the/a/an* [PATIENT argument] […]
A verb found within the verbs of creation and transformation which describes a type of "performance". This performance is in itself the affected object. Accordingly, anything that is a result of the named action can occupy the argument slot (e.g., *book*).

4.2 The LexTALE stimuli

The vocabulary test consisted of 60 items (40 words and 20 nonwords), in addition to 3 practice items (see Appendix 7).[18] These were selected from a pool of 240 items based on a pilot study (Lemhöfer & Broersma, 2012) where participants made a word/nonword decision on all 240 items. Separately for the words and the nonwords, four categories of difficulty were formed according to the percentage of correct scores. This transpired in the following manner: for each stimulus, the item-whole correlation was calculated to validate how well an item discriminates good from poor participant performance. The 25% with the highest item-whole correlation in each of the four categories were included in the vocabulary test.

All the items of the LexTALE are between 4 to 12 letters long (mean: 7.3). The 40 words have a mean frequency of 1 to 26 (mean: 6.3) ipm according to the CELEX online lexical database (http://celex.mpi.nl/). The words are composed of 15 nouns, 12 adjectives, 1 verb, 2

[18] The three practice items act as dummies and are not taken into account for the calculation of the score.

verb participles, 2 adverbs, in addition to 8 words that belong to two different syntactic classes (e.g., both a noun and a verb, such as *dispatch*). All the real words followed British English rather than American English spelling. The nonwords are orthographically and pronounceably fit nonsense strings of letters created by either changing some letters in an already existing word (e.g., *proom*) or by recombining existing morphemes (e.g., *rebondicate*). Lastly, the order of items was fixed so that no more than five words or nonwords would appear in a row.

5 Analysis

5.1 Analyzing the performance on the sentence completion test

The primary data consisted of having participants fill in the blanks for incomplete sentences. To evaluate the performance of the L2 users, a control group assuming the role of the L1 population was necessary. In line with previous research of similar nature that utilized native speakers to juxtapose L2 data with L1 data (see Johnsen, 2016; Reine, 2016) the present study did so too – though in a different manner. In addition to the participant group, three corpora datasets of English were included. Conjointly, these served as a representation of native speaker responses and acted as the target response (or *reference standard*) against which participants were compared. More explicitly, the primary response provided by the L2 participants as a completion to a given sentence trial was set side by side with the most frequently occurring response in the corpus data.

When deciding on where the focus of interest should be located regarding participant responses, what part of the performance data was important for the mapping of L2 sensitivity to argument structure had to be determined. Ultimately, three aspects attracted particular attention. The most obvious being the extent to which participants overlapped with native speakers. In addition, it was deemed just as important to see how many responses that are similar to the target response (i.e., a response that contains the same semantic features as the target), and how many that do not match it (i.e., a response that does not correspond to the target response). The prediction was that more proficient speakers would be more likely to generate the target response of a given stimulus due to advantages related to (meta)linguistic competence. In addition, it was expected to find more incorrect responses in less proficient speakers. These predictions are consistent with how proficiency and sentence-level semantic processing co-operate with verb constraint in determining NP-filler choice. An increase in proficiency leads to a higher chance of selecting a more accurate NP-filler for a given verb – provided that the verb poses restrictions on the internal argument slot.

To analyze the performance data, what constitutes a unique response had to be clarified. When judging and counting responses, only NP heads were counted. The same was the case for number (*singular, plural*) and articles (*the, a, an*). Although the latter remark may seem to contradict the instruction in the sentence completion test to always start a response with an article, this was only to ensure that participants' responses followed the same syntactic structure used in the corpus searches (i.e., VERB – ARTICLE – NOUN). As such, *Easter egg, an egg,* and *the eggs* would all count as the same response – *egg*.

One limitation of the corpus searches is that the results only express the syntax command, and nothing more. The use of the syntactic structure generated only a single noun following the article. To achieve compatibility with the corpus data, an NP composed of a longer string of words would be abbreviated into a singular, equivalent NP. Accordingly, a response consisting of (e.g.,) a book title would be boiled down to represent its essence – *The Lord of the Rings* would correspond to *book*. Furthermore, responses like *telly, television*, and *TV*, would all be counted as the same unique response – *television*. Plus, any spelling errors were corrected to match the intended word (*chair* instead of **chear*), and the number of blank responses per verb was tracked (cases in which some participants did not respond).

The aspect of sense was not taken into consideration. Although some NPs have sets of multiple meanings the word may denote, there was no way of knowing which sense the participant referred to. For example, in the sentence 'S*he wrote*…', a possible NP-filler might be *a letter*. However, whether the intended meaning was for the NP to denote a written message, or a written symbol that represents a sound of speech is irrelevant. As such, the intended meaning of a response did not matter as long as it satisfied the selectional restrictions set by the verb.

5.2 Analyzing the performance on the vocabulary test

The nature of YES/NO decision tests enables several possible methods to evaluate performance – and LexTALE is no exception. In a validation study of the vocabulary test, Lemhöfer and Broersma (2012) administered the test to a group of participants, then compared participants' performance across a set of scoring protocols. In the end, the scoring method that provided the highest correlation between score and proficiency was *% correct$_{AV}$* (*average percentage correct*)[19]. This measure is calculated by averaging the proportion of correct responses for words and nonwords. However, since LexTALE consists of an unequal amount of correct vs incorrect items (40 words and 20 nonwords), simply calculating the average of all items would yield a bias as a rejection of a real word would be penalized more than the acceptance of a nonword. The *% correct$_{AV}$* scoring protocol accounts for this lack of symmetry by purely averaging the percentage of correct responses for words and nonwords.

---

[19] ((number of words correct/40*100) + (number of nonwords correct/20*100)) / 2

5.3 Statistical treatment

The raw data collected by participants across both online tests was filled into Microsoft Excel and imported into the statistical analysis software *IBM SPSS 27*[20] which treated all the data analytics.

[20] Statistical Package for the Social Sciences (version 27) by International Business Machines Corporation.

6 Results

The following chapter will present the core findings of the study as derived from the method applied to both gather and analyze information.

6.1 Performance on the vocabulary test

Based on the performance of participants on the vocabulary test, general L2 proficiency could be calculated in accordance with the *Common European Framework of Reference for Languages* (CEFRL). This framework of language competence describes language learner's ability in terms of reading, listening, speaking, and writing. The six reference levels used in this regard are portrayed in Figure 1.

Figure 1

The CEFRL levels

Proficient user	C2 | Mastery C1 | Effective Operational Proficiency
Independent user	B2 | Vantage B1 | Threshold
Basic user	A2 | Waystage A1 | Breakthrough

The CEFRL covers two main dimensions: a vertical and a horizontal one. The vertical dimension is divided into three broad divisions ranging from basic to proficient user (displayed with the letters A to C). The horizontal dimension divides each proficiency level into two sublevels to separate between lower and upper users of the respective group (displayed with the number 1 and 2).

To display LexTALE scores as CEFRL proficiency levels, the vocabulary scores needed to be converted into correlating general competence levels. Lemhöfer and Broersma (2012) calculated which ranges of LexTALE scores are associated with which CEFRL levels. Their estimate did, however, not discriminate between the C levels. This element was accounted for in the present study to capture the division among the proficient users (see Table 3).

Table 3

Correlation between CEFRL proficiency levels and LexTALE scores

CEFRL level	CEFRL description	LexTALE score
C2	Upper proficient user	91-100%
C1	Lower proficient user	81-90%
B2	Upper independent user	61-80%
B1 and lower	Lower independent user	≤ 60%

With all this in mind, Table 4 shows the participants' scores on the vocabulary test. Of the 54 participants, only 1 (1.8%) achieved CEFRL level B2 with a LexTALE score of 60%, whereas the rest managed to reach B1 or higher. Thus, all participants were placed within the proficiency range of B and C level. It is interesting to note that over half of the informants (51.9%) achieved C2 (the highest rank). The findings indicate that as far as vocabulary knowledge is concerned, the participants performed well above average. Seeing that the vocabulary test served as a proxy for general language proficiency of English, it is safe to assume that the participants of this study are highly proficient L2 users (M = 86.8; SD = 11.2).

Table 4

Distribution of participants across CEFRL levels

CEFRL level	No. of participants	Percentage
C2	28	51.9%
C1	10	18.5%
B2	15	27.8%
B1 and lower	1	1.8%
TOTAL	54	100%

6.2 Performance on the sentence completion test

The following segment will consider the result of the sentence completion test on grounds of several perspectives. First and foremost, the general performance of participants (*en masse*) will be highlighted with reference to their distribution of responses for every verb. As stated in the Analysis (Section 5.1), three aspects were of interest: 1) the hit rate accuracy to the target response; 2) the number of similar responses; and 3) the number of incorrect responses.

6.2.1 Performance as a group

Table 5

Mean values for the total number of unique responses per verb in each verb category

		Variable of constraint		
		Most constrained	Moderately constrained	Least constrained
Variable of frequency	Frequent	15.30	19.10	22.50
	Infrequent	11.60	16.60	28.10
	Averaged	13.45	17.85	25.30

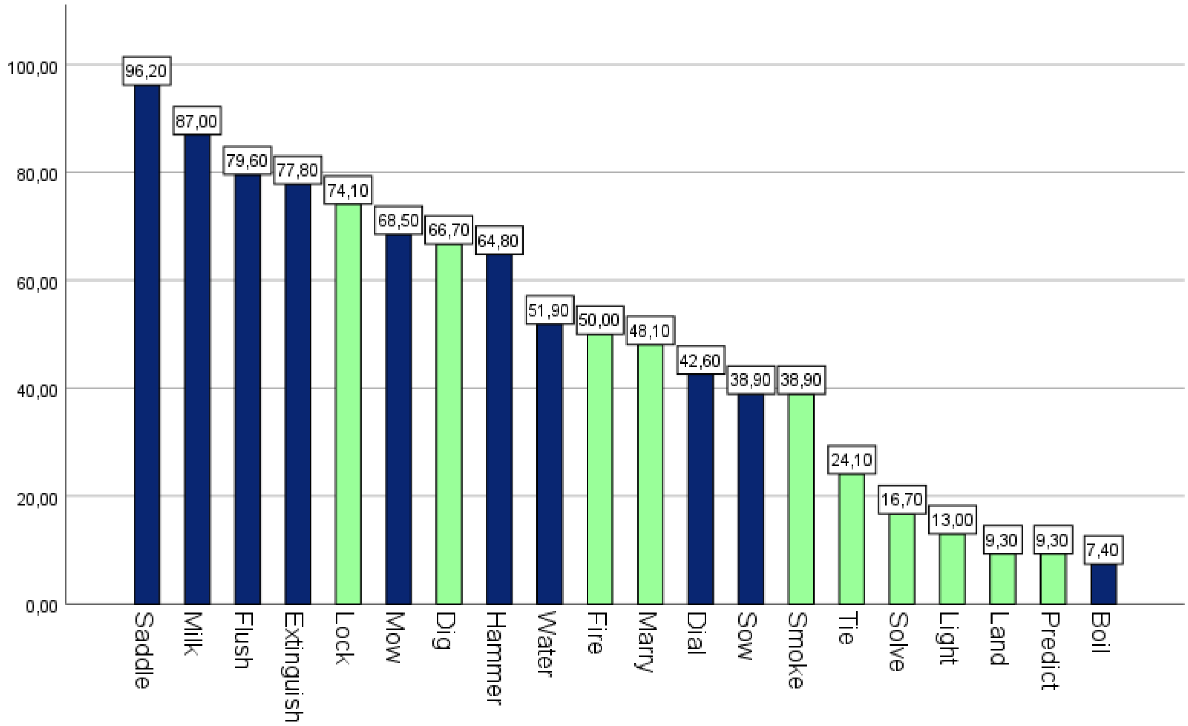
Table 5 gives an overview of mean values in each of the verb categories. Here the interest is located at the total number of unique responses provided per verb on average by the participating group. As hypothesized, there is a relationship between the degree of constraint and the amount of received unique responses described in the following terms: parallelly as the level of constraint decreases, the total number of unique responses reported per verb increases – and vice versa. On the other hand, and contrary to expectations, it appears that the infrequent verbs (in the most- and moderately constrained categories) acquired less unique responses on average, opposed to the frequent verbs. Although a frequency effect of some sort was anticipated, this was not the envisioned one. Both tendencies will be addressed further in the discussion (Section 7.4).

6.2.1.1 Performance on the most constrained verbs

Figure 2 presents the relationships between the primary responses and the target responses for the most constrained verbs. For every verb, L2 performance was compared against that of the corpora databases (representing the L1 population). The target NP-filler response for a given verb was determined by its frequency of occurrence in the corpora: the most accustomed NP-filler associated with a verb was the one that made up the preponderance portion of the total number of NPs.

Figure 2

Percentage of participant responses that matched the target responses for the most constrained verbs



Note: Verbs of light color are frequent verbs; verbs of dark color are infrequent verbs.

Table 6 gives a complete summary of how participants responded to the most constrained verbs. For each verb in the chart, the following information is provided: 1) the target response (set by the corpora); 2) the primary response (being the NP-filler that accumulated the highest amount of the total responses); 3) the percentage of similar responses (responses that did not match the target response but still fulfilled the selectional requirements); 4) the percentage of incorrect responses (responses that did neither match the target response nor fulfilled the selectional requirements); and 5) the percentage of blank responses.

Table 6

Distribution of responses to the most constrained verbs

Most constrained category	Target response(s)	Participant performance				
		Primary response		Similar responses	Incorrect responses	Blank responses
Saddle	horse	horse	96.2%	1.9%		1.9%
Milk	cow	cow	87.0%	8.5%	4.5%	
Flush	toilet	toilet	79.6%		18.5%	1.9%
Extinguish	fire	fire	77.8%	7.6%	14.6%	
Lock	door	door	74.1%	16.4%	7.6%	1.9%
Mow	lawn	lawn	68.5%	18.2%	13.3%	
Dig	hole	hole	66.7%	16.8%	14.6%	1.9%
Hammer	nail	nail	64.8%	14.3%	20.9%	
Boil	water | kettle	egg	61.1%	38.9%		
Land	job	plane	53.7%	33.0%	13.3%	
Water	plant	plant	51.9%	42.4%	5.7%	
Fire	gun	gun	50.0%	36.7%	13.3%	
Marry	woman | man	woman	48.1%	31.0%	19.0%	1.9%
Dial	number	number	42.6%	55.5%		1.9%
Sow	seed	seed	38.9%	28.3%	32.8%	
Smoke	cigarette	cigarette	38.9%	59.2%	1.9%	
Solve	problem	puzzle	31.5%	64.7%	3.8%	
Light	fire	candle	27.8%	60.8%	11.4%	
Tie	knot/game	knot	24.1%	75.9%		
Predict	future	outcome	13.0%	84.2%	3.8%	

Note: Some verbs have two or three target responses in the target response column. This is owing to incoherence between the relevant corpora (iWEB, COCA, and BNC) in terms of what NP-filler expressing the highest frequency to follow a specific verb. This tendency occurs at a higher rate for the moderately- and (especially) the least constrained verb category.

Despite the strict sets of selectional requirements the most constrained verbs pose, every primary response fulfills its respective verb's semantic criteria. This trend is reflected in two definable aspects of the collected data. One is the large portion of the total responses being located in the primary response- and the similar response category; while the other is in the relatively few responses being found in the incorrect response- and the blank response category. In addition to responding adequately (by selecting NP-fillers that hold the pertinent semantic properties for the target verbs) participants also produced, for the most part, the target responses. The only exceptions of target response and primary response mismatches were *solve, light, land, predict,* and *boil*. Interestingly, these five verbs make up the tail end of Figure 2:

solve problem with 16.7%, *light* fire with 13.0%, *land* job with 9.3%, *predict* future with 9.3%, and *boil* water with 7.4%.

Another point to remark is the distribution of incorrect responses. Although the main bulk of responses per verb is classified as adequate responses, exactly half of the verbs display a high number of incorrect responses – having an incorrect response value of 10% or higher. These are *light* (11.4%), *fire* (13.3%), *land* (13.3%), *mow* (13.3%), *extinguish* (14.6%), *dig* (14.6%), *flush* (18.5%), *marry* (19.0%), *hammer* (20.9%), *sow* (32.8%). Most of these verbs have incorrect response values between the 10 to 20% mark. There is, however, one verb that sticks out from this norm, *sow*. Considering the strict selectional requirements set by the verb, responses provided to this verb (and to that of *mow*) will be discussed further in Section 7.3.

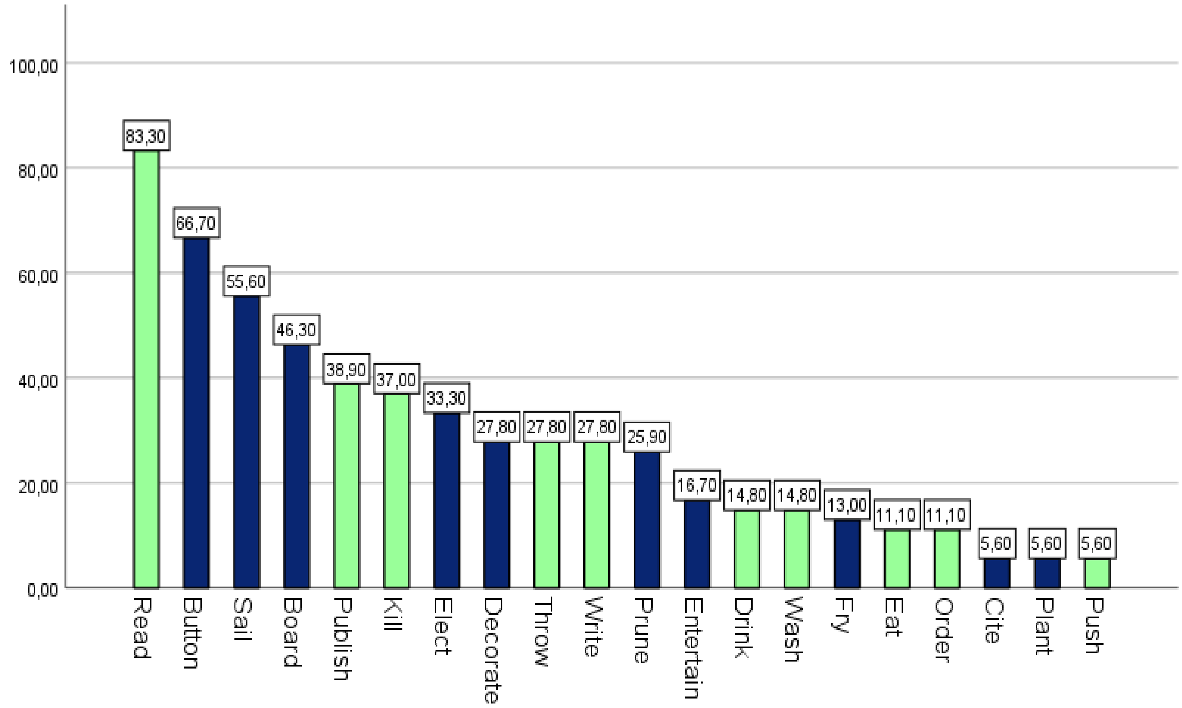
Furthermore, some sort of the frequency effect alluded to in Table 5 is again evident. The NP-filler responses to infrequent verbs are more unison than responses to frequent verbs. This bias towards the less occurring verbs is apparent given the fact that infrequent verbs are situated at the left side of Figure 2. (Bear in mind that the verbs are ordered after the primary response value which means that participants match the target response of native speakers exceedingly more often and to a greater degree for the infrequent verbs opposed to the frequent ones.) Another observance that might be the consequence of the frequency variable is the fact that the five mismatch cases were all frequent verbs, with *boil* being the only exception.

6.2.1.2 Performance on the moderately constrained verbs

Figure 3 presents the relationships between the primary responses and the target responses for the moderately constrained verbs. When compared to the similar bar chart for the most constrained verbs (Figure 2), the present one is less supplementary at all levels. Primary responses to the moderately constrained verbs make up a significantly weaker portion of the total responses; failing to reach the same height percentagewise. Rather, the primary responses are located at the lower end of the scale by a significant margin.

Figure 3

Percentage of participant responses that matched the target responses for the moderately constrained verbs



Note: Verbs of light color are frequent verbs; verbs of dark color are infrequent verbs.

As demonstrated in Table 7 below, there is once more a common occurrence for primary responses and target responses to correlate with one another. As expected, however, this tendency comes into play to a lesser degree in contrast to the most constrained verbs – a result that is most likely a cause of the reduced sets of selectional requirements enforced by the moderately constrained verbs. This decrease in constraint is also witnessed by the corpora. Similar to the participants, the corpora were also in agreement with one another to a lower extent. The multiple target responses for most verbs testify to the differing results between the corpora in this regard. In total, seven verbs exhibited a lack of overlaps. This was the case for *plant, drink, fry, cite, push, wash*, and *eat*. Again, verbs subjected to primary response and target response mismatches made up the lower region of Figure 3: *wash* <u>*dishes*</u> received 14.8%, *drink* <u>*water*</u> with 14.8%, *fry* <u>*egg*</u> with 13.0%, *eat* <u>*food*</u> with 11.1%, *cite* <u>*source*</u> with 5.6%, *plant* <u>*seed*</u> with 5.6%, and *push* <u>*button*</u> with 5.6%.

Table 7

Distribution of responses to the moderately constrained verbs

Moderately constrained category	Target response(s)	Participant performance				
		Primary response		Similar responses	Incorrect responses	Blank responses
Read	book	book	83.3%	16.7%		
Button	shirt | jacket	shirt	66.7%	33.3%		
Plant	seed	tree	55.6%	33.0%	11.4%	
Sail	sea/boat	boat	55.6%	42.5%	1.9%	
Board	plane | train	plane	46.3%	46.2%	7.5%	
Publish	book	book	38.9%	61.1%		
Kill	man | enemy	man	37.0%	59.2%	3.8	
Elect	president	president	33.3%	42.1%	24.6%	
Drink	water	beer	29.6%	62.8%	7.6%	
Decorate	house | room | wall	house	27.8%	68.4%	3.8%	
Throw	ball	ball	27.8%	66.6%	5.6%	
Write	book | letter	book	27.8%	70.3%	1.9%	
Prune	tree | bush	bush	25.9%	43.9%	24.6%	5.6%
Fry	onion | egg	chicken	25.9%	74.1%		
Cite	source	book	25.9%	72.2%	1.9%	
Push	button	cart	24.1%	74.0%	1.9%	
Wash	dishes | car	house	18.5%	79.6%	1.9%	
Entertain	crowd | audience | kid	crowd	16.7%	46.1%	37.2%	
Eat	food	banana	14.8%	85.2%		
Order	book | drink | pizza	pizza	11.1%	87.0%		1.9%

Another observation of Table 7 is the relatively few responses located in the incorrect category. This can be seen as a verification that the amount of possible NP-fillers for the moderately constrained verbs is larger than that of the constrained verbs. The smaller range of selectional requirements put onto the argument slot allows for more entities to fulfill the criteria; a trend that leads to a diminished chance of inserting an inadequate NP. There are only four outliers that show an incorrect response value greater than 10%. These are *plant* (11.4%), *elect* (24.6%), *prune* (24.6%), and *entertain* (37.2%). Of which, *elect* and *prune* will be discussed more thoroughly later as they demonstrate L1 transfer (Section 7.2) and a deficient lexical representation (Section 7.3), respectively.

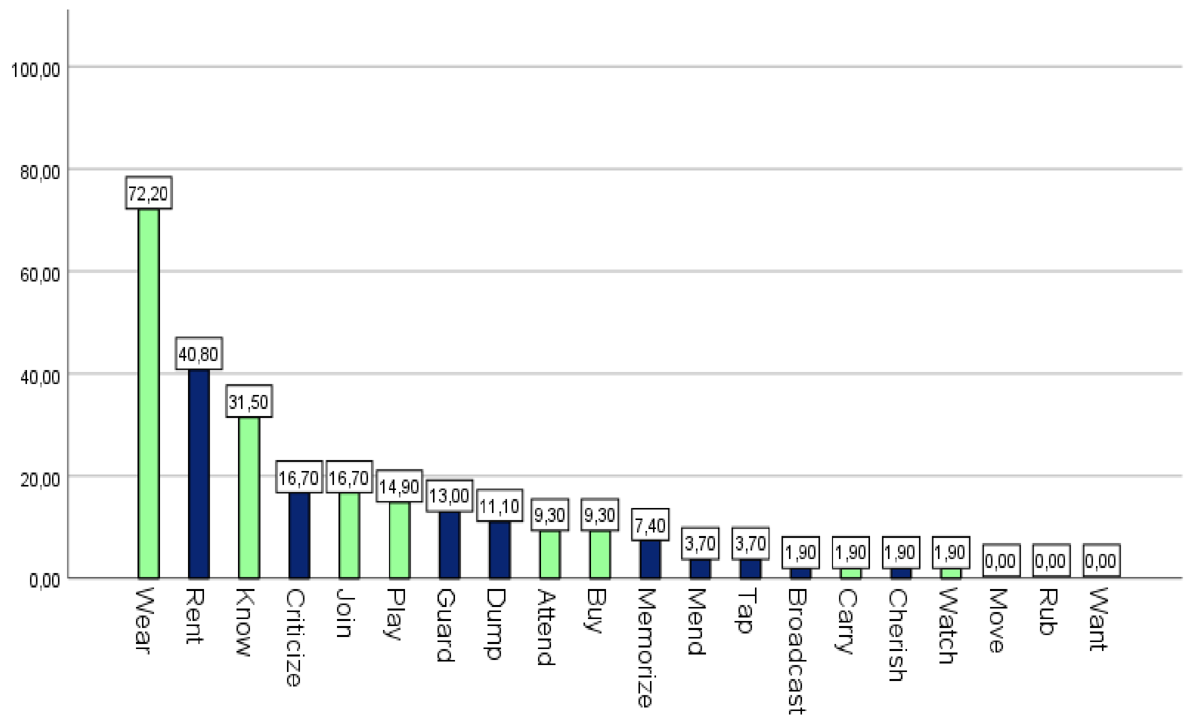
The influence of frequency is reaffirmed once again to be biased towards infrequent verbs. For the moderately constrained verbs, there is a correlation between frequency and overlaps between primary response and target response. Infrequent verbs show a larger degree of overlaps between the L1 and L2 population with most of the mismatches occurring for frequent verbs.

6.2.1.3 Performance on the least constrained verbs

Although the least constrained verbs represent the baseline verbs, Figure 4 and Table 8 are provided to show that the trends witnessed from the most- to moderately constrained verbs will continue from the moderately- to the least constrained verbs.

Figure 4

Percentage of participant responses that matched the target responses for the least constrained verbs



Note: Verbs of light color are frequent verbs; verbs of dark color are infrequent verbs.

In line with expectations, the performance on the least constrained verbs was inferior compared to that of the previous verb categories. Even though matchings between primary response and target response do occur, Figure 4 is evidence that it happens at a low ratio. Apart from *wear*, *rent*, and *know*, the remaining least constrained verbs received overlaps below 20% with four verbs receiving only 1 response (*broadcast* 1.9%, *carry* 1.9%, *cherish* 1.9%, *watch* 1.9%) and three receiving none whatsoever (*mow* 0.0%, *rub* 0.0%, *want* 0.0%).

Table 8

Distribution of responses to the least constrained verbs

Least constrained category	Target response(s)	Participant performance				
		Primary response		Similar responses	Incorrect responses	Blank responses
Wear	helmet | hat | dress	dress	72.2%	25.9%	1.9%	
Move	cursor | ball	car	44,4%	55.6%		
Rent	car | house	car	40.8%	59.2%		
Dump	water | body	girl	33.3%	76.7%		
Know	answer	answer	31.5%	66.6%		1.9%
Tap	button | spacebar | screen	table	25.9%	72.2%		1.9%
Buy	house	apple	24.1%	75.9%		
Attend	event | meeting	class	22.2%	75.9%	1.9%	
Broadcast	game | tournament	show	22.2%	77.8%		
Play	game	piano	22.2%	77.8%		
Join	team | club | conversation	army	20.4%	77.7%		1.9%
Rub	skin | ball	lamp	20.4%	79.6%		
Cherish	moment | environment	man	18.5%	81.5%		
Watch	video | match	movie	18.5%	79.6%		1.9%
Criticize	government | president	government	16.7%	81.4%		1.9%
Want	job	car	16.7%	83.3%		
Carry	gun	bag	13.0%	87.0%		
Guard	entrance | house | door	door	13.0%	75.1%	1.9%	
Memorize	word | path | list	book	11.1%	88.9%		
Mend	relationship | fence	wound	11.1%	88.9%		

Table 8 displays only five cases where primary responses match target responses, being *wear*, *rent*, *know*, *guard*, and *criticize*. Appropriately, these verbs are situated at the far left of

Figure 4 with *wear* having 72.2%, *rent* with 40.8%, *know* with 31.5%, *criticize* with 16.7%, and *guard* with 13.0%. Undoubtedly, this limited portion of overlaps is a consequence of the openness of the least constrained verbs. Nevertheless, there is a trade-off of bearing few selectional requirements. Although exhibiting an open nature entails there is no genuine way of predicting the target response, it follows that almost any entity can fill the argument slot. This unrestrictive character implies that there is hardly any NP that cannot succeed such a verb. It is precisely this proneness to fillers that is seen in the nearly empty column of incorrect responses with only three incorrect responses overall (*wear* 1.9%, *attend* 1.9%, *guard* 1.9%).

For the least constrained verbs, no considerable difference between high- and low frequency verbs was observed. The only polarity is the existence of more unique responses provided for the infrequent ones. Regardless, it seems that the concept of verb frequency/familiarity does not play a role in principally unconstrained verbs.

6.2.2 Correlation between level of performance and level of proficiency

Considering that the preceding Figures and Tables of prior segments only captured the performance of participants as a whole, the following section will look more closely at the performance of participants based on their L2 proficiency.

To determine the relationship between the L2 proficiency and L2 performance, Pearson's correlation coefficients (*r*) were calculated. This is a statistical measure of the strength and direction of the linear relationship between a pair of variables. By design, it is constrained to take a value in the fixed range -1 to +1 ($-1 \leq r \leq +1$), within which the strength of a relationship must be. The stronger the correlation, the closer the correlation coefficient is to either extreme (±1). On the other hand, the closer to zero, the less strength there is in the correlation at play. Table 9 shows the proposed guidelines by Cohen (1988) for how to interpret the strength of a coefficient correlation.

Table 9

Strength of association

	Coefficient, *r*	
Strength of association	Positive	Negative
Small	.1 to .3	-.1 to -.3
Medium	.3 to .5	-.3 to -.5
High	.5 to 1.0	-.5 to -1.0

To decide whether the linear relationships observed in the sample data can be used to model the relationship of the population, the probability value (p-value) must be considered in relation to *the null hypothesis* (H_0) and *the alternative hypothesis* (H_1).[21] In this regard, the p-value is used in significance testing to evaluate whether the obtained data supports ($p \neq 0$) or rejects ($p =0$) H_0. In other words, it represents the probability that the correlation between variable *x* and *y* occurred by chance. The threshold value to what is considered statistically significant, and thus reject H_0, is a p-value of 5% ($\alpha = 0.05$). This indicates that the risk of concluding that a correlation exists – when, in fact, no correlation exists – is 5%. Therefore, if the p-value is smaller than the set significance level ($p \leq 0.05$), it can be concluded without reasonable doubt that the correlation between two variables is statistically significant. In this event, H_0 is rejected in favor of H_1. If, on the other hand, the p-value is larger than the significant level ($p > 0.05$), there is not strong enough evidence to suggest an effect exists in the population. In this case, H_0 holds true.

This being said, the correlation between the results of the vocabulary test and the sentence completion test is reported in Table 10.

[21] These are contradictory with the following opposing viewpoints:

i. H_0: There is no relationship between the two variables.
ii. H_1: There is a relationship between the two variables.

Table 10

Correlation between LexTALE scores and performance on each verb category

		LexTALE score	Most constrained category	Moderately constrained category	Least constrained category
LexTALE score	Pearson Correlation	1	.341(*)	.020	-.118
	Sig. (2-tailed)		.012	.885	.397
	N	54	54	54	54
Most constrained category	Pearson Correlation	.341(*)	1	.261	-.010
	Sig. (2-tailed)	.012		.057	.941
	N	54	54	54	54
Moderate constrained category	Pearson Correlation	.020	.261	1	.135
	Sig. (2-tailed)	.885	.057		.329
	N	54	54	54	54
Least constrained category	Pearson Correlation	-.118	-.010	.135	1
	Sig. (2-tailed)	.397	.941	.329	
	N	54	54	54	54

Note: *Correlation is significant at the 0.05 level (2-tailed).

Table 10 shows that there is a significant correlation between vocabulary knowledge and the sentence completion test for the most constrained category (p = 0.012 < 0.05). This is a positive and moderate relationship with a correlation coefficient of 0.341. There was observed no relationship between the vocabulary test and the performance on the moderately constrained category with a correlation coefficient of 0.020; p = 0.885 > 0.05. The same is also the case for the vocabulary test and the least constrained category with no significance correlation (p = 0.397 > 0.05) and a correlation coefficient of -0.118. Accordingly, it appears that proficiency bears only a deterministic role in sentences of highly constrained contexts.

A similar conclusion can be drawn when considering the number of target response matches per individual in each of the verb categories (see Appendices 8, 9, 10). For the most constrained category, individuals of the B2 group provided on average 8.73 target response overlaps (SD: 1.86). Those of the C1 group provided 10.10 target response overlaps (SD: 2.55), whereas participants of the C2 group provided 10.32 target response overlaps (SD: 2.35). In this case, there is a clear distinction between the B2 group[22] and the C groups – the higher the proficiency, the better performance. However, as constraint diminishes so does this correlation between proficiency and performance. For the moderately constrained category, the B2 participants produced on average 7.33 target response overlaps (SD: 2.09); the C1 participants with 6.50 target response overlaps (SD: 0.97); and the C2 participants with 7.03 target response overlaps (SD: 1.75). For the least constrained category, the B2 groups provided 2.93 target response overlaps (SD: 1.27) on average; the C1 group with 2.70 target response overlaps (SD: 1.15); and the C2 group with 2.85 target response overlaps (SD:1.55). Except for the most constrained category, the bottom proficiency group performed the most like the L1 population, statistically speaking. Still, the rate of performance across proficiency groups was much the same for the lesser constrained verb categories. As Table 10 shows, the level of proficiency is only reflected in the most constrained category.

[22] The sample size of the B1 group was too small to average (*n* = 1) and was therefore excluded.

7 Discussion

7.1 Sensitivity and proficiency

Verbs denote events. As such a verb provides the most information to a sentence. Its lexical entry stores semantic- and syntactic information that decides how its argument structure shall unfold. However, semantics and syntax do not play equal parts in this context; rather, the syntactic expression of a verb's argument structure is defined by its semantic side (being specified by information on participants). The lexicalization of a participant's syntactic category and semantic role is therefore assumed to be pre-determined by the semantic properties of the verb. This gives rise to a predictive relationship between a verb and its arguments where information stored in the verb's lexical entry can be exploited to predict upcoming post-predicate argument(s). Such lexically encoded information is however only accessible given sufficient vocabulary knowledge (by the speaker). Accordingly, the idea underlying this thesis is the existence of a link between vocabulary knowledge and lexical semantics of verbs in which a learner's sensitivity to argument structure can be estimated on the grounds of her level of vocabulary knowledge.

The conducted lexical decision test (LexTALE), an experimental way of assessing participants' reaction to 'words', acted as a measure to assess L2 vocabulary knowledge. Considering the importance of vocabulary knowledge to understand and communicate in a foreign language, the estimated vocabulary knowledge of a participant has been judged a good predictor of her general proficiency in the target language. Studies on the matter prove this clearly, showing a moderate to strong relationship between vocabulary knowledge and each basic language ability (read, write, listen, and speak) (e.g., Milton et al., 2010). Given this close connection, vocabulary knowledge is commonly associated with CEFRL language levels; as one's score in a vocabulary test can be used as a reliable placement measure. As previously presented, Table 4 shows that participants were located at the uppermost portion of the English competence scale (mean rank of lower proficient users; C1).

Given the overall high L2 proficiency of participants, it is only fitting that the results from the sentence completion test show that the participants respond in a native-like manner to a large degree. In general, responses fulfill the selectional requirements set by the given verb; thereby matching (or being similar as) the target response. Still, even for a participating group consisting of advanced L2 users, the influence of proficiency is evident. Participants with higher vocabulary scores were the ones that also, on a general basis, matched the target responses more frequently and to a larger extent. But this held only true in the most constrained condition (otherwise, proficiency proved to be insignificant). Based on the correlation coefficient measures in Table 10, the only statistically significant correlation is between proficiency and the most constrained category ($p < 0.05$). This positive-going relationship implies that an increase in one variable yields an increase in the other: the higher the proficiency, the more target response overlaps. The same result is not repeated for the other categories with no significant correlation for the moderately constrained category ($p > 0.05$) and the least constrained category ($p > 0.05$). This is in line with previous research on the matter. Johnsen (2016) and Reine (2016) found evidence of participants' gaze proportions being dependent on their level of proficiency: the higher proficiency, the more looks towards the target item. Identical to the result from the present project, however, their findings suggest that this phenomenon is exclusive for sentences where the verb was of a highly restrictive semantical character.

On the basis of this report, and contrary to expectations, it appears that proficiency (in the form of vocabulary knowledge) impacts primarily cases where choice is highly constrained. The influence of proficiency bears a less deterministic role for the less semantically strict verbs; rather, it prospers in the narrower lexical network with limited collocational alternatives. This indicates that in the highly constraining sentence contexts, in which but a few completion candidates are adequate to follow the verb, participants with high proficiency are more likely to respond in the exact same manner as native speakers. However, the farther away from a highly constraining sentence, the less proficiency matters.

All things considered, the findings suggest that every participant, regardless of proficiency level, expresses some form of sensitivity to argument structure. This favors a gradient view of sensitivity opposing that of Johnsen (2016) who hypothesized that C1 marks the proficiency level threshold "at which sensitivity to argument structure is developed enough to employ it in a predictive manner in language processing" (p. 30). Still, the degree of this sensitivity manifestation is dependent on proficiency: the closer a learner embodies a native-like behavior, the greater sensitivity is displayed.

7.2 L1 transfer

The advanced L2 proficiency of participants is affirmed in the high accuracy rate between primary responses and target responses. Still, overlaps are not observed in all situations. Responses to certain verbs reveal that participants have treated the verbs somewhat differently from the native English-speaking population. Upon closer inspection, it seems that this tendency is rooted in the L1 influence on the L2.

By nature, humans possess complete native-like mastery in their L1 – an achievement thought (by many) to be impossible in an L2. Accordingly, the former is more entrenched in the mind of a speaker; being a crucial part of one's thinking. At this point, recall Stringer's (2010) *lexical transfer* argument of how L1 lexical semantics constitutes the initial state of L2 lexical acquisition (as noted in Section 2.5.2). When first subjected to SLA, learners are prone to resort to their pre-existing L1 knowledge of lexical items when treating L2 items. This is an intuitive reaction simply because of the need to lean on some form of underlying linguistic knowledge. Although this strategy might work in the most general sense, it will not account for the broad range of subtle differences associated with a shift from one lexicon to another. However, L1 transfer occurs primarily at the early stages of SLA and decreases as proficiency increases. Whether learners can reach a level where they are not subjects to transfer is another question for another paper.

The following subsections will take a closer look at verbs that seemed to have been perceived through L1 lenses and try to understand why these evoked transfer from the L1 lexicon. An overview of the unique responses for the verbs to be discussed can be found in Appendix 11.

7.2.1 Boil

For the incomplete sentence *"Roy boiled …"* a native English speaker will, according to the corpora, most likely respond with either *water* (iWeb & COCA) or *kettle* (BNC). The results of the test, on the other hand, show that the participants seldom if ever matched either one of the target responses, with *water* paired 7.4% and *kettle* 0%. Instead, *egg* was their preferred response by a large margin (61.1%). This begs the question of why the target responses did not invite any attention.

To respond to the matter, it is worth drawing attention to the selectional requirements set by the verb. *Boil* denotes the preparation of food which may either occur 1) by causing the temperature to reach a degree at which a liquid, edible entity bubbles, or 2) by cooking a solid, edible entity in boiling water. The verb has two different senses that allow for two different approaches to which it can be realized. By the look of it, it seems the different treatment of *boil* by the sample populations is rooted in which sense of the verb that holds the strongest activation in each respective group. The favored sense will bring about the corresponding property of the edible entity (liquid versus solid). Given that the samples represent the respective populations at large gave birth to the following hypothesis:

> For boil, an L1 speaker of English prefers an NP-filler denoting a liquid to follow given the stronger entrenchment of the first sense. On the other hand, an L1 speaker of Norwegian prefers an NP-filler denoting a solid given the stronger entrenchment of the second sense.

To make sure that this suspicion carried weight, another sentence completion test was conducted. Following the same exact rules and form as the initial test (Appendix 6), this second sentence completion test was designed for native speakers of Norwegian on Norwegian verbs (see Appendix 12). The test revealed that the verb *koke* 'boil', in the sentence "*Eirik kokte ...*", was once again followed most frequently by *egg* 'egg' (53.8%). Interestingly, when analyzing the remaining responses, these carried also the property of solid food: *potet* 'potato' (23.1%), *grønnsak* 'vegetable' (7.7%), *knoke* 'knuckle' (7.7%), *ris* 'rice' (7.7%). The responses correlate with those elicited from the English sentence completion test on two fronts. First of all, the primary response to each sentence completion test overlapped. Secondly, responses to each test were generally NPs bearing the property of solid food. Unfortunately, due to time constraints, a similar sentence completion test was not conducted for native speakers of English on English verbs. In light of this further lack of evidence on the side of the English-speaking population, a more credible hypothesis on the matter would instead be:

> For boil, an L1 speaker of Norwegian prefers an NP-filler denoting a solid given the stronger entrenchment of the second sense.

Be that as it may, the treatment of *boil* appears to be a consequence of lexical transfer. The second sentence completion test reveals that Norwegian holds a stronger activation to solid food for *koke*. This tendency seems to translate over to the corresponding L2 verb, *boil*.

7.2.2 Land

Another verb that indicates L1 influence on L2 is shown in response to *land*, presented in the incomplete sentence *"Moe landed ..."*. Given the underlying syntactical expression of the predicate (*land the/a/an* NOUN), the most frequent NP-filler in each corpus was *job*. This was, however, not the primary response of the participants who only matched the target response with 7.4%. Alternatively, *plane* was their most frequent response (53.7%).

With its selectional requirements, the most common usage of the verb is to describe either 1) someone's success in achieving something desirable, or 2) the act of bringing an aircraft to the ground. Based on this description of *land*, it looks as though each sample population again favors one sense of the verb over the other. From the perspective of a Norwegian, the preference for *plane* is self-explanatory when considering the limits of the equivalent Norwegian verb *lande* 'land'. The dictionary *Bokmålsordboka*[23] (https://ordbok.uib.no/) provides two definitions for the verb. The first describes the act of bringing an aircraft to the ground; the other the act of bringing something ashore. Hence one is shared with the English language, the other is not. Thus, the languages differ in terms of what events the verb conceptualizes. Whereas (25) shows the usage of land possible in both languages, (26) does not.

(25)

a. He	**landed**	a plane.
b. 'Han	**landet**	et fly.'

(26)

a. He	**landed**	a job.
b. *'Han	**landet**	en jobb.'
c. 'Han	**fikk**	en jobb.'

The examples above demonstrate how Norwegian uses two separate verbs to express two separate happenings, while English needs only utilize two separate senses of one verb to express two separate happenings. This difference is witnessed in (26b) where maintaining the verb *å lande* 'to land' gives rise to a questionable sentence for a semantic-pragmatic reason. Instead, the verb *å få* 'to get' must be employed as in (26c).

Going back to the sentence completion test, responses seem to reflect the limits of the L1 verb. With responses such as *spaceship*, *rocket*, *helicopter*, and so on, it appears that the L1 representation of the verb manifests a stronger entrenchment. As a result, few responses go beyond the confines of the L1 verb.

[23] An online, monolingual dictionary for modern *Bokmål* produced in a cooperative project between the University of Oslo and the Norwegian Language Council.

7.2.3 Elect

The unique responses to the verb *elect* in the incomplete sentence "*He elected …*" show also signs of transfer. This tendency seems to be the cause of no one-to-one lexical equivalence of *elect* in Norwegian; rather, the verb carries multiple possible translations.

(27)

a. He	**elected**	…
b. 'Han	**besluttet**	…'
c. 'Han	**bestemte**	…'
d. 'Han	**kåret**	…'
e. 'Han	**utpekte**	…'
f. 'Han	**valgte**	…'
g. Etc.		

Participants who matched the responses of the L1 population perceived the verb to bear the message "to choose a candidate in an election"; hence a translation into an appoint verb such as 'å kåre' in (27d). This reading of the verb embodies the selectional restrictions of a human entity, specifically for political office. The results indicate that the majority of participants provided NP-fillers that abide by these semantic features by either overlapping with the target response (*president*) or providing responses of similar nature (e.g., *candidate* or *politician*). Still, there were a significant number of responses that did not meet the selectional requirements of the verb (24.6%). These responses appear to be rooted in some participants having a more general reading of the verb ("to make a choice from a number of alternatives") resulting in a translation into 'å velge' as in (27f). While this verb denotes the same meaning as *to elect* in one sense, it corresponds to the meaning of *to select* in another. Importantly, 'å velge' encompasses a broader field of use allowing for two separate events to be expressed through two unique senses of the verb. In contrast, two separate verbs must be employed in English to perform the same conceptions.

(28)

a. 'Han	**valgte**	en kandidat.'
b. He	**elected**	a candidate.

(29)

a. 'Han	**valgte**	et nummer.'
b. ?? He	**elected**	a number.

c. He **selected** a number.

Accordingly, responses like (e.g.) *kiwi, puppy,* and *sofa* following *to elect* imply that some participants have treated *elect* with *select* in mind. This trend is likely due to the stronger activation between 'å velge' and the reading "to make a choice from a number of alternatives" than it is between the verb and the reading "to choose a candidate in an election" in Norwegian. Essentially, this showcases the L1 influence on the L2 as the verb holds a stronger association with *select* than *elect* during translation due to its most common usage in Norwegian.

7.3 Deficient lexical representation

As mentioned in the theory chapter, exposure to target language input has been identified as a critical component of SLA. However, linguistic units differ in their frequency of occurrence which entails inequality of items' exposures. According to the *lexical quality hypothesis* (Perfetti & Hart, 2002), all words are stored in the mind as lexical representations in relation to the three well-integrated defining constituents of orthography, phonology, and semantics. (The lexical representation of the word *cat* has the orthographical portrayal C-A-T, the phonological depiction /kæt/, and the semantic representation of whatever *cat* means.) All the same, words are not stored in equal status and they vary in the quality of their lexical representations; a phenomenon thought to be linked to frequency as it is a determining factor that influences the strength in which a word is entrenched in the mind. A high frequency word is prone to a high-quality representation due to its higher rate of recurrence in the input which increases the chance of developing a familiarity with its orthographic-, phonologic-, and semantic information. In contrast, the opposite is the case for a low frequency word.

The following subsections will present three verbs several participants had a faulty perception of. An overview of the unique responses for the verbs to be discussed can be found in Appendix 13.

7.3.1 Sow

Across the three tables showing the distribution of participant responses to each verb category (Tables 6,7,8), the verb *sow* received one of the highest numbers of incorrect responses (32.8%). Interestingly, for the sentence '*Irwin sowed …*', all incorrect responses were of similar nature (i.e., entities of fabric), including NP-fillers such as *blanket* (5.6%), *dress* (5.6%), *sweater* (3.7%), and so on. In view of this, it appears that several participants have instead of processing the provided verb *sow*, activated the homophonic competitor *sew*. These respective verbs have shared properties of orthographic- and phonological form. Firstly, they exchange a similar appearance: each word is three letters long with an identical onset and coda syllable structure, only differing at the nucleus. Secondly, each word has the phonological representation /səʊd/ in the past tense. This makes the words stored in our minds in relation to each other at two representational levels. In other words, their lexical entries are interconnected. With orthography and phonology into account, when encountering the target verb *sow*, the rivaling homophonic verb *sew* was activated (along with its different meaning). Thus, if subject to having a low-quality representation of *sow*, it might cause the rivaling word to be preferred, whereas having a high-quality representation will presumably only cause a brief interference.

I hypothesize the relevant participants to have a low-quality representation of *sow*. This deficient orthographical representation – and not a phonological one seeing that these are indistinguishable – is what is assumed to have led to the activation of the wrong word during reading.

7.3.2 Mow

The semi-homophonic pair *mow*/*move* was also observed. In reaction to the sentence fragment, '*Curtis mowed …*', responses such as *chair* (5.6%), *couch* (1.9%), *truck* (1.9%), and so on were found. This suggests that some participants processed the verb *move* over *mow*. This is probably due to shared properties of orthographic- and partly phonological form. In the past tense, the words are quite identical with the only spelling difference is between the use of '*w*' and '*v*': M-O-W-E-D versus M-O-V-E-D. As for their pronunciations, these differ somewhat more noticeably: *mowed* /məʊd/ versus *moved* /muːvd/.

Based on previous research, the activation of *move* rather than *mow* was most likely a result of the frequency differences between the two words. It is assumed that reading the more frequent word of a homophonic pair does not yield an activation of the less frequent one (Hogaboam & Perfetti, 1978). Following this line of thought, the more experience with a word, the stronger its orthographic-, phonologic-, and semantical information is stored. This leads to a more stable lexical unit compared with a less frequent word. Going back to the present incomplete sentence, I hypothesize that the activation of *move* was a combined effect of orthography and frequency. This entails a deficient orthographic representation of *mow* which triggered the semi-homophonic competitor *move*. And since the contender word is a highly frequent verb whose lexical quality is in all probability high, it trumps the infrequent verb with its low-quality lexical representation. In other words, *move* has a faster lexical retrieval from the lexical storage due to its stronger entrenchment as a result of its high frequency of occurrence, opposed to the low frequency word *mow*.

7.3.3 Prune

The verb *prune* also suffered from having a low-quality lexical representation by several participants. However, unlike the previous examples that indicated a deficient orthographic representation, the present reveals a deficient semantic representation. The incomplete sentence '*Karen pruned …*' elicited a large number of incorrect responses (24.6%) with no particular cohesion; e.g., *friendship* (1.9%), *database* (1.9%), *thought* (1.9%), and so on. In addition to the verb receiving the most blank responses (5.6%), it appears that many participants did not know its meaning.

This indicates an unreliable, impoverished representation of the word. Thus, whether a participant is familiar with its pronunciation and spelling is ruined by the lack of word meaning. This deficiency thwarts the threefold mapping between the word's orthographic-, phonologic-, and semantic form. Accordingly, what word meaning participants had in mind when responding is unknown – except that it was not the target one. The insufficient familiarity with the word is in all likelihood an outcome of its infrequency: limited exposure increases the probability of establishing a low-quality representation with regards to its semantic aspect.

7.4 The influence of constraint and frequency

The initial assumption that different properties of verbs yield differing argument structure realization seems correct. The variables of constraint and frequency, from which a network of six verb categories was developed (see Table 1), demonstrate this effect in a gradient and relatively consistent manner. However, whereas the variable of constraint performed as expected, the same was not true for that of frequency.

7.4.1 Consistent with expectations

The stimuli sentences used in this experiment were designed to differ in constraint at three levels, identified as the most-, moderately-, and least constrained (verb) categories. A given sentence's degree of constraint was determined by the selectional requirements the verb of that sentence posed onto the following argument slot. In this sense, being highly constraining entails having a large set of selectional requirements. This tendency is reflected in the more lexical-semantic demands an NP-filler candidate must fulfill to even be considered selected as the argument. By contrast, being less constraining presupposes few selectional requirements which, in turn, permits a wide range of NPs to be potential NP-filler candidates. With this in mind, the nature of the three categories can be viewed as a three-level pyramid where each verb category represents one level on the pyramid: starting with the least constrained category at the bottom, it becomes narrower the farther up the pyramid we move.

Furthermore, the concept of constraint is also tied to that of predictability – as previously noted in Section 2.4. A highly constraining verb helps built up the contextual environment in which the sentence the verb is part of takes place to a larger degree than less constraining verbs. The more contextual information provided, the easier is the process of predicting upcoming input. In this sense, a highly constraining verb establishes a high-predictability environment; a moderately constraining verb establishes a moderately-predictability environment; and a least constraining verb establishes a minimal-predictability environment.

The results of the sentence completion test indicate that the information encoded in the verbs about their selectional requirements and contextual information was available and contributed to predictable effects. However, this phenomenon was dependent on the level of constraint the verb embodies. The most constrained verbs received more target response overlaps and less unique responses per verb. Whereas the farther away a verb is from being highly constraining, the less target response overlaps and more unique responses are observed. This demonstrates the differences between the verb categories efficiently. The stricter nature of the most constrained verbs guides participants into an appropriate post-predicate filler by its narrower semantic constraint. This is accompanied by the fact that only a limited pool of NP-fillers can follow the verb in the first place. On the contrary, the less constraining verb categories are less strict as to what NP-filler represents the argument. The more open property of these verbs allows for more potential NP-filler candidates to be selected which lowers the chances of responding in the exact same manner as native speakers. This is especially true for the least constrained verbs in which there is an almost non-existing filtering method in the form of selectional requirements the verbs place onto the argument slot. This allows for a vast body of NPs to be selected which makes the idea of participants answering in a synchronized manner nothing but a remote illusion.

7.4.2 Contrary to expectations

The frequency with which linguistic tokens occur in the input has been established as a crucial component in SLA. With the presumption that argument structure is situated in the verb, its reoccurrence rate is of heightened interest as greater frequency leads to an enhanced verb knowledge of form-meaning mapping. This stance is keeping with the idea that knowledge of linguistic constructions depends on experience of use (Larsen-Freeman, 1994; MacWhinney, 1999).

Based on this correlation between (verb) frequency and (verb) knowledge, the original hypothesis was that the variable of frequency would favor highly occurring verbs in the sentence completion test. Comparatively, a frequent verb was expected to be treated in a unison manner to a larger degree than an infrequent verb. The former is (most likely) more familiar to the speaker which, in turn, suggests that what NP-filler usually follows is ingrained in the mind.[24] This was, however, not the case. Contrary to the hypothesized association between frequency and the distribution of responses, the infrequent verbs outperformed the frequent ones in the most- and moderately constrained category. In general, participants' performance on infrequent verbs accumulated the least unique responses (being more unison) and matched the target responses to a higher degree percentagewise. At the same time, the frequent verbs witnessed more mismatches.

With reference to the abovementioned observations, two potential explanations are presented to resolve the findings. The first is linked to verb familiarity. More experience with a verb unlocks a greater arsenal of argument slot filler candidates the learner may utilize to succeed the verb. This increase in potential fillers leads to a decrease in the chance of a homogeneous response. On the other side, less experience with a verb results in a weakened familiarity which limits the supply of NP-fillers adequate to follow the verb – at least that is known to the user. Due to this small pool of possible competitors to choose from (relative to that of the more familiar verbs), there is a reinforced probability to respond in a unison manner. The second possible explanation is rooted in the idea that infrequent verbs have more distinct characteristics than frequent verbs. Because of extensive prior exposures, a high frequency verb is expected to have links to several linguistic contexts (as in settings the verb might be situated in). Conversely, in return for being less experienced, a low occurring verb exhibits fewer links

[24] Section 3.1.2 on 'verb frequency': […] the expectation was that frequent verbs would generate few unique responses due to them being more known and thus encourage uniform answers; whereas low occurring verbs would promote an increase in unique responses due to their lesser recognized nature.

and instead manifests a more unique representation. A by-product of this profile is that there are fewer NP-fillers associated with the internal argument slot of the verb.

Regardless of the cause, the results indicate a frequency effect that appears to hold a stronger position than anticipated when it comes to its influence regarding what NP-filler is to follow a verb depending on the frequency quality of this verb.

7.5 Importance of the findings

I find it interesting, but maybe even more so enticing, that my initial thoughts on what the effects of frequency would be, failed. Not only was I incorrect but the complete opposite result occurred. Seeing that I had based my assumption on previous research regarding linguistic frequency effects, could it simply be an odd, irregular outcome – one that has no reason of occurring other than by sheer coincidence – we are witnessing? Of course. Do I however think so? No.

The differences in elicited responses to verbs of an infrequent nature contra those of a frequent one are too prominent for such an "easy way out"-explanation. Just as the level of constraint has already been recognized as a deterministic component in narrowing the pool of NP-filler candidates down, I believe the level of frequency makes up the other component. Still, it has its limits.

Figure 5

Hierarchy of the variables

Constraint

▼

Frequency

The results indicate that constraint holds a more prominent influence than frequency. This is clear when considering how the effect of constraint is eminent across the board while the frequency effect is only significant in the most constrained condition. Hence, constraint acts independently of frequency but not the other way around. What we end up with is a view of frequency as an enhancement effect to the variable of constraint, only evident in highly constraining contexts – an effect that decreases in tandem with that of constraint.

7.6 Limitations of the study

The study is not without its shortcomings. The design of the stimuli sentences is one such flaw. The decision to include an adverbial phrase (AdvP) as a default syntactic function at the end of every sentence, and to have the subject always be an AGENT represented by either a proper noun (e.g., *Shelly*) or a personal pronoun (e.g., *He*)[25] was done to introduce some variance between the sentences. Everyone who has volunteered for an academic study has most likely experienced some form of boredom. Monotonous, unengaging, and uninteresting (or simply put *bland*) stimuli nurse such boredom. You can feel your attention span slowly but steadily decreases as you feel the study cannot come to an end fast enough. As a result, you put less effort into your role in the study to *get it over with* as fast as possible. The inclusion of AdvPs and shifting subjects (for a lack of a better word) served to counteract this. There was a tradeoff with this, however. It added more linguistic units which generated more lexical information – context. And since learners make use of all information available in a sentence when anticipating what items will come next (as previously mentioned by Matsuki et al., 2011), this might reduce the predictive effects of the verbs' lexical entries. In hindsight, excluding AdvPs and opting for a gender-neutral, semantically minimal pronoun (e.g., *someone*) might have been the right call after all. It would at least have isolated the predictive effects of a verb to a greater degree.

Another weakness of the study is the verb categorization. Although it worked, it is far from optimal. Listings of such kind are hard. You do not have any standardized method to follow and scarce research to draw from. But worst of all, you do not know how your itemization will work out until the study has concluded – and at that point, it is too late. Yes, there was a clear-cut division in how participants responded to these, but there were still some verbs that should have been removed in favor of another, potentially more successful verb. One that would have reflected its respective category more accurately.

[25] Here I am referring to every sentence starting off with a unique subject, instead of having the same subject across all sentences.

8 Conclusion

The aim of the study was to investigate to what extent native speakers of Norwegian are sensitive to English argument structure, and the influence of verb constraint and verb frequency in this regard. As addressed in my analysis of recent research, no evidence suggests that some participants did *not* display sensitivity. Instead, every participant displayed some form of sensitivity, regardless of one's proficiency. From this point of view, the concept of sensitivity exhibits gradience. Still, in line with previous research on the matter, proficiency is influential but only in highly constraining contexts. The expectation that differing L2 proficiency levels yielding differing L2 sensitivity levels were not fully met. The findings suggest that a learner's proficiency level holds only a deterministic role in the most constraint conditions, whereas proficiency remained inconsequential in the others. In this sense, a learner's level of proficiency can in turn explain her level of sensitivity to argument structure to highly constraining verbs.

The expectations regarding the influence of the verb properties in responses to the incomplete sentences were not entirely fulfilled either. On the one side, the property of constraint impacted as envisaged. The more constraint the context provided by the verb became, the more target response overlaps per verb were elicited (most > moderately > least). The property of frequency, on the other side, did not influence as initially hypothesis. Instead of the high frequency verbs evoking more target response overlaps, the discovered frequency effect revolved around the potency of low frequency verbs to do so (low > high).

As mentioned in the review of the literature, studies investigating L2 learners' sensitivity to argument structure are scarce. Prior studies are quite recent and need to be supplemented with further empirical research. This study is one step in that direction. Still, we have but merely scratched the surface of the topic and there remain many gaps in our knowledge around L2 sensitivity to argument structure. Based on the unexpected findings of the frequency effect in the present study, this highlights a new avenue that could be explored in future studies. Given that this frequency effect is true (and not a one-off), it would be interesting to see whether the frequency effect is gradient similar to how the constraint effect is gradient. Establishing three (or more) groups of varying levels of frequency would be necessary to explore how responses are influenced the less frequent a verb becomes.

Work cited

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

Baker, M. (1988). *Incorporation: a theory of grammatical function changing*. University of Chicago Press.

Borer, H. (2005). *Structuring Sense*. Oxford University Press.

Bowerman, M., & Brown, P. (2008). *Crosslinguistic perspectives on argument structure*. Taylor and Francis Group.

Bresnan, J. (1995, October 12, 1995). *Lexicality and Argument Structure* Syntax and Semantics Conference, Paris.

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

Butt, M. (2008). *Theories of Case*. Cambridge University Press.

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

Chomsky, N. (1981). *Lectures on Government and Binding*. Foris Publications.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates.

De Saussure, F. (1959). *Course in General Linguistics*. New York, Philiosophical library.

Delong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from brain activity. *Nature neuroscience*, *8*(8), 1117-1121.

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

Dimitriova-Vulchanova, M., & Weisgerber, M. (2007). 'Concept', 'Context' and beyond: manner of motion verbs, 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.

Ellis, N. C. (2002). Frequency effects in language acquisition: a review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, *24*, 143-188.

Enayat, J., Amirian, M., Zareian, S. M. R., & Ghaniabadi, S. G. (2018). Reliable Measure of Written Receptive Vocabulary Size: Using the L2 Depth of Vocabulary Knowledge as a Yardstick. *SAGE Open*, *8*(1), 2158244017752221. https://doi.org/10.1177/2158244017752221

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning & Verbal Behavior*, *12*(6), 627-635. https://doi.org/10.1016/S0022-5371(73)80042-8

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

Frisch, S., Hahne, A., & Friederici, A. D. (2004). Word category and verb-argument structure information in the dynamics of parsing. *Cognition*, *91*(2004), 191-219.

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Affectedness and direct objects: the role of lexical semantics in the acquisition of verb argument structure. *Cognition*, *41*(1), 153-195.

Gyllstad, H. (2013). Looking at L2 Vocabulary Knowledge Dimensions from an Assessment Perspective – Challenges and Potential Solutions. In C. Bardel, B. Laufer, & C. Lindqvist (Eds.), *L2 vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis* (Vol. Eurosla Monographs Series, 2, pp. 11-28). EUROSLA. http://eurosla.org/monographs/EM02/getfile.php?name=Gyllstad

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

Hamberger, M. J., Friedman, D., & Rosen, J. (1996). Completion norms collected from younger and older adults for 198 sentence contexts. *Behavior Research Methods, Instruments and Computers*, *28*, 102-108.

Hammerås, M. L. (2017). *Probing sensitivity to argument structure in two proficient level groups - an exploratory study with Norwegian learners of English* NTNU, Dragvoll]. Trondheim.

Hellan, L., & Dimitriova-Vulchanova, M. (2000). Criteriality and grammatical realization. In S. Coopmans, T. Everaert, & J. Grimshaw (Eds.), *Lexical specification and insertion*. John Benjamins.

Hersen, M. (2003). *Comprehensive handbook of psychological assessment volume 2: personality assessment*. John Wiley and Sons.

Hogaboam, T. W., & Perfetti, C. A. (1978). Reading skill and the role of verbal experience in decoding. *Journal of Educational Psychology*, *70*(5), 717-729.

Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. The MIT Press.

Jackendoff, R. (2002). *Foundations of Language*. Oxford University Press.

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

Juffs, A. (2000). An overview of the second language acquisition of links between verb semantics and morpho-syntax. In J. Archibald (Ed.), *Second Language Acquisition and Linguistic Theory* (pp. 197-227). Blackwell Publishing.

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

Katz, J. J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language*, *39*(2), 170-210.

Kiss, T. (2011). Reconstituting Grammar: Hagit Borer's exoskeletal syntax. A review article. *Linguistic Berichte*.

Koenig, J. P., Mauner, G., & Bienvenue, B. (2003). Arguments for adjuncts. *Cognition*, *89*(2), 67-103. https://doi.org/10.1016/S0010-0277(03)00082-9

Langacker, R. W. (1987). *Foundations of cognitive grammar : 1 : Theoretical prerequisites* (Vol. 1). Stanford University Press.

Larsen-Freeman, D. (1994). Chaos/complexity science and second language acquisition. *Applied Linguistics*, *18*, 141-165.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: a quick and valid Lexical Test for Advanced Learners of English. *Behav Res Methods*, *44*(2), 325-343. https://doi.org/10.3758/s13428-011-0146-0

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

Levin, B., & Rappaport Hovav, M. (1998). Building verb meanings. In M. Butt & W. Geuder (Eds.), *The projection of arguments: lexical and compositional factors* (pp. 97-134). CSLI Publications.

Levin, B., & Rappaport Hovav, M. (2005). *Argument realization*. Cambridge University Press.

MacWhinney, B. (1999). *The emergence of language*. Erlbaum.

Matsuki, K., Chow, T., Hare, M., Elman, J. L., Scheepers, C., & McRae, K. (2011). Event-based plausibility immediately influences on-line language comprehension. *Journal of experimental psychology: learning, memory, and cognition*, *37*, 913-934.

Milton, J., Wade, J., & Hopkins, N. (2010). Insights into Non-native Vocabulary Teaching and Learning. In C.-B. Rubén, A.-C. Christian, & T.-L. María del Mar (Eds.), *Chapter 6. Aural Word Recognition and Oral Competence in English as a Foreign Language* (pp. 83-98). Multilingual Matters. https://doi.org/doi:10.21832/9781847692900-007

Murray, H. A., & MacKinnon, D. W. (1946). Assessment of OSS personnel. *Journal of Consulting Psychology*, *10*, 76-80.

Naigles, L. G., Gleitman, H., & Gleitman, L. R. (1993). Children acquire word meaning components from syntactic evidence. In E. Dromi (Ed.), *Language and Cognition: a developmental perspective* (pp. 104-140). Ablex.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*(4), 273-281. https://doi.org/10.1080/17470216508416445

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden path sentences: evidence of the application of verb information during parsing. *Journal of experimental psychology: learning, memory, and cognition*, *20*, 786-803.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (Vol. 11). John Benjamin.

Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge: MIT Press.

Reine, A. S. (2016). *L2 speakers' argument structure sensitivity inferring from the participants' eye gazes: An eye-tracking experimental study with adult Norwegians learners of English* NTNU, Dragvoll]. Trondheim.

Rotter, J. B. (1951). Word association and sentence completion methods. In H. H. Anderson & G. L. Anderson (Eds.), *An introduction to projection techniques* (pp. 279-310).

Saeed, J. I. (2016). *Semantics* (Fourth edition. ed., Vol. 2). Wiley Blackwell.

Schachter, J. (1993). A new account of language transfer. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 98-111). John Benjamins.

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and Full Transfer / Full Access model. *Second Language Research*, *12*(1), 40-72.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*, 109-147.

Sprouse, R. A. (2006). Full transfer and relexification: L2 acquisition and creole genesis. In C. Lefebvre, L. White, & C. Jourdan (Eds.), *L2 acquisition and creole genesis: Dialogues* (pp. 169-181). John Bejamins.

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

Stringer, D. (2019). Lexical semantics: relativity and transfer. In N. Erdogan & M. Wei (Eds.), *Applied linguistics for teachers of culturally and linguistically diverse learners* (pp. 180-203). IGI Global.

Tendler, A. D. (1930). A preliminary report on a test for emotional insight. *Journal of Applied Psychology*, *14*, 122-136.

Tenny, C. L. (1994). *Aspectual roles and the syntax-semantics interface*. Kluwer.

Tolkien, J. R. R. (2005). *The Lord of the Rings* (63 ed.). HarperCollins*Publishers*.

Treffers-Daller, J., & Calude, A. (2015). The role of statistical learning in the acquisition of motion event construal in a second language. *International Journal of Bilingual Education and Bilingualism*, *18*(5), 602-623.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: evidence from ERPs and reading times. *Journal of experimental psychology: learning, memory, and cognition*, *31*(3), 443-467.

Van Patten, B., & Benati, A. G. (2010). *Key terms in second language acquisition*. Continuum Internation Publishing Group.

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*(2012), 176-190.

Wassenaar, M., & Hagoort, P. (2007). Thematic role assignment in patients with Broca's aphasia: sentence-picture matching electrified. *Neuropsychologia*, *45*, 716-740.

Williams, E. (1981). Argument structure and morphology. *The Linguistic Review*, *1*, 81-114.

Wright, C. E. (1979). Duration differences between rare and common words and their implications for the interpretation of word frequency effects. *Memory & Cognition*, *7*(6), 411-419.

Yue, Y., & Fan, S. (2016). *Measurement of Vocabulary Knowledge*. Rotterdam: Sense Publishers. https://doi.org/10.1007/978-94-6300-524-1_16

Åfarli, T. (2007). Do verbs have argument structure? In E. J. Reuland, T. Bhattacharya, & G. Spathas (Eds.), *Argument Structure* (Vol. 108). Amsterdam: John Benjamins Publication.

Appendix 1: Raw and normalized frequency

		Frequency (raw)			Frequency (ipm)		
		Corpora			Corpora		
No.	Verb	iWeb	BNC	COCA	iWeb	BNC	COCA
1	Attend	696 749	3 497	30 717	49.76	34.97	30.71
2	Board	82 596	434	4 279	5.89	4.34	4.27
3	Boil	86 237	438	5 005	6.15	4.38	5.00
4	Broadcast	129 453	824	8 723	9.24	8.24	8.72
5	Button	60 939	123	1 052	4.35	1.23	1.05
6	Buy	2 706 584	11 929	143 673	193.32	119.29	143.67
7	Carry	1 114 476	9 823	65 452	79.61	98.23	65.45
8	Cherish	38 662	141	2 983	2.76	1.41	2.98
9	Cite	58 629	274	6 910	4.18	2.74	6.91
10	Criticize	37 825	299	7 721	2.70	2.99	7.72
11	Decorate	92 903	360	3 021	6.63	3.60	3.02
12	Dial	69 457	194	3 427	4.96	1.94	3.42
13	Dig	212 034	878	16 957	15.14	8.78	16.95
14	Drink	508 015	2 981	39 913	36.28	29.81	39.91
15	Dump	77 979	287	6 992	5.56	2.87	6.99
16	Eat	1 442 144	7 225	113 436	103.01	72.25	113.43
17	Elect	71 555	410	6 210	5.11	4.10	6.21
18	Entertain	84 297	636	5 793	6.02	6.36	5.79
19	Extinguish	13 795	92	945	0.98	0.92	0.94
20	Fire	187 961	731	19 678	13.42	7.31	19.67
21	Flush	49 790	177	2 809	3.55	1.77	2.80
22	Fry	56 570	419	3 267	4.04	4.19	3.26
23	Guard	96 058	542	9 960	6.86	5.42	9.96
24	Hammer	18 625	154	1 676	1.33	1.54	1.67
25	Join	1 343 693	7 094	65 330	95.97	70.94	65.33
26	Kill	817 223	4 239	128 100	58.37	42.39	128.10
27	Know	12 514 493	117 874	2 110 687	893.89	1 178.74	2 110.68
28	Land	184 467	1 225	13 840	13.17	12.25	13.84
29	Light	178 071	841	11 561	12.71	8.41	11.56
30	Lock	156 118	612	11 241	11.15	6.12	11.24
31	Marry	144 725	2 531	26 707	10.33	25.31	26.70
32	Memorize	34 142	55	2 112	2.43	0.55	2.11
33	Mend	12 101	264	1 465	0.86	2.64	1.46
34	Milk	4 255	31	392	0.30	0.31	0.39
35	Move	2 238 168	13 064	176 939	159.86	130.64	176.93
36	Mow	18 738	76	1 446	1.33	0.76	1.44
37	Order	549 205	1 832	21 773	39.22	18.32	21.77
38	Plant	148 962	719	9 154	10.64	7.19	9.15
39	Play	3 142 135	14 093	200 932	224.43	140.93	200.93
40	Predict	202 562	1 328	19 204	14.46	13.28	19.20

		Frequency (raw)			Frequency (ipm)		
		Corpora			Corpora		
No.	Verb	iWeb	BNC	COCA	iWeb	BNC	COCA
41	Prune	19 786	103	650	1.41	1.03	0.65
42	Publish	255 254	1 294	11 957	18.23	12.94	11.95
43	Push	682 378	2 891	48 029	48.74	28.91	48.02
44	Read	4 620 366	16 575	277 562	330.02	165.75	277.56
45	Rent	237 541	654	9 456	16.96	6.54	9.45
46	Rub	119 671	594	7 473	8.54	5.94	7.47
47	Saddle	5 077	48	641	0.36	0.48	0.64
48	Sail	90 475	879	4 693	6.46	8.79	4.69
49	Smoke	143 731	1 087	14 718	10.26	10.87	14.71
50	Solve	511 075	1 886	29 941	36.50	18.86	29.94
51	Sow	31 609	188	1 567	2.25	1.88	1.56
52	Tap	266 991	767	9 082	19.07	7.67	9.08
53	Throw	575 107	2 899	55 233	41.07	28.99	55.23
54	Tie	199 955	988	13 736	14.28	9.88	13.73
55	Want	11 360 506	54 141	1 081 589	811.46	541.41	1 081.58
56	Wash	292 170	1 606	16 133	20.86	16.06	16.13
57	Watch	1 551 577	6 212	135 159	110.82	62.12	135.15
58	Water	24 652	71	1 000	1.76	0.71	1.00
59	Wear	1 013 984	4 142	55 707	72.42	41.42	55.70
60	Write	1 682 925	10 497	115 273	120.20	104.97	115.27

Appendix 2: The verb categories

Most constraint and frequent verb category	Target response in each corpus					
	iWeb		BNC		COCA	
Tie	game	18.94%	knot	20.26%	knot	14.72%
Lock	door	17.45%	door	50.37%	door	63.44%
Light	fire	13.85%	fire	26.20%	fire	18.11%
Fire	gun	5.74%	gun	14.70%	gun	7.81%
Smoke	cigarette	17.43%	cigarette	24.95%	cigarette	33.08%
Dig	hole	31.00%	hole	27.82%	hole	33.41%
Solve	problem	51.05%	problem	60.89%	problem	53.13%
Land	job	20.88%	job	11.66%	job	15.56%
Predict	future	20.15%	future	6.66%	future	17.07%
Marry	man	12.23%	man	15.02%	man	17.23%

Most constraint and infrequent verb category	Target response in each corpus					
	iWeb		BNC		COCA	
Extinguish	fire	33.50%	fire	32.14%	fire	19.06%
Milk	cow	32.59%	cow	25.00%	cow	36.02%
Hammer	nail	15.14%	nail	16.66%	nail	14.18%
Mow	lawn	62.42%	lawn	76.92%	lawn	72.70%
Water	plant	26.04%	plant	25.00%	plant	24.02%
Sow	seed	71.08%	seed	85.71%	seed	73.18%
Dial	number	27.37%	number	45.23%	number	37.43%
Saddle	horse	17.93%	horse	57.14%	horse	32.75%
Boil	water	17.08%	kettle	25.75%	water	16.66%
Flush	toilet	21.43%	toilet	15.78%	toilet	37.43%

Moderately constraint and frequent verb category	Target response in each corpus					
	iWeb		BNC		COCA	
Read	book	10.47%	book	9.26%	book	13.36%
Drink	water	8.52%	water	9.81%	water	9.83%
Eat	food	3.74%	food	4.87%	food	3.63%
Write	book	6.64%	letter	10.26%	book	8.09%
Publish	book	7.33%	book	7.31%	book	3.34%
Wash	car	4.07%	dish	8.02%	dish	13.07%
Throw	ball	8.84%	ball	7.00%	ball	13.41%
Order	book	1.66%	drink	1.39%	pizza	4.69%
Kill	enemy	3.08%	man	8.02%	man	5.75%
Push	button	6.25%	button	2.63%	button	7.39%

Moderately constraint and infrequent verb category	Target response in each corpus					
	iWeb		BNC		COCA	
Fry	onion	10.56%	onion	50.00%	egg	15.67%
Plant	seed	19.78%	seed	15.21%	seed	20.00%
Board	plane	14.38%	train	13.33%	plane	22.88%
Decorate	house	3.62%	room	9.47%	wall	7.78%
Sail	sea	18.98%	boat	20.68%	boat	18.24%
Prune	tree	15.51%	bush	8.44%	tree	17.50%
Entertain	kid	7.64%	crowd	4.46%	audience	3.79%
Button	shirt	6.07%	shirt	25.00%	jacket	6.52%
Cite	source	9.14%	source	4.34%	source	5.36%
Elect	president	13.03%	president	3.07%	president	29.95%

Least constraint and frequent verb category	Target response in each corpus					
	iWeb		BNC		COCA	
Buy	house	4.34%	house	4.85%	house	6.31%
Know	answer	5.39%	answer	5.41%	answer	5.29%
Wear	helmet	3.95%	hat	3.19%	dress	3.34%
Watch	video	3.95%	match	1.18%	video	5.18%
Move	cursor	2.56%	cursor	5.23%	ball	4.78%
Play	game	20.84%	game	11.69%	game	18.84%
Carry	gun	2.02%	gun	1.33%	gun	6.06%
Join	team	5.30%	club	3.12%	conversation	10.94%
Want	job	1.14%	job	1.32%	job	1.75%
Attend	event	5.77%	meeting	12.58%	meeting	5.91%

Least constraint and infrequent verb category	Target response in each corpus					
	iWeb		BNC		COCA	
Broadcast	game	4.19%	tournament	1.66%	game	4.00%
Cherish	moment	6.74%	environment	5.55%	moment	9.17%
Criticize	government	5.76%	government	5.17%	president	11.39%
Dump	water	3.74%	body	11.11%	body	10.72%
Guard	entrance	3.50%	house	3.29%	door	4.69%
Memorize	word	2.69%	path	14.28%	list	4.41%
Rent	car	25.14%	house	11.57%	car	18.88%
Rub	skin	2.03%	ball	7.40%	skin	2.52%
Tap	button	3.68%	spacebar	10.30%	screen	2.37%
Mend	relationship	7.99%	fence	6.97%	fence	8.33%

Appendix 3: Information sheet

Forespørsel om deltakelse i forskningsprosjektet

«Andrespråksbrukeres prosessering av engelske setninger»

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å undersøke hvordan norske fremmedspråkbrukere av engelsk prosesserer engelske setninger. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Bakgrunn og formål.
Jeg er en masterstudent ved Institutt for språk og litteratur ved NTNU som jobber med et forskningsprosjekt der jeg trenger universitetsstudenter til å delta. Prosjektet har fokus på tilegnelse av andrespråk, og målet med studiet er å undersøke hvordan norske fremmedspråkbrukere av engelsk prosesserer engelske setninger.

Hvem er ansvarlig for forskningsprosjektet?
Institutt for språk og litteratur ved NTNU er ansvarlig for prosjektet.

Hvorfor får du spørsmål om å delta?
Enhver universitetsstudent med norsk som morsmål vil i utgangspunktet ha mulighet til å delta i studiet.

Hva innebærer deltakelse i studien?
Deltakerne vil gjennomføre et elektronisk spørreskjema i løpet av høsthalvåret 2020. Spørreskjemaet vil inneholde to språktester. Først vil deltakerne bli presentert med en rekke ufullstendige engelske setninger som skal fullføres. Deretter vil deltakere utføre en vokabulartest der man skal bestemme om diverse enkeltord er en del av det engelske språket eller ikke.

Frivillig deltakelse.
Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykket tilbake uten å oppgi noen grunn. Alle dine personopplysninger vil da bli slettet. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger.
Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrivet. Alle personopplysninger vil bli behandlet konfidensielt og i samsvar med personvernregelverket. Dette vil si at den enkeltes opplysninger vil bli erstattet med en kode som lagres på egen navneliste adskilt fra øvrige data. Det vil kun være veileder som har tilgang til listen som knytter navn til deltakernummer, og denne vil lagres utilgjengelig for uvedkommende. Enkeltpersoner vil ikke kunne gjenkjennes i publikasjonen.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?
Prosjektet skal etter planen avsluttes i juni 2021, og personopplysninger vil da slettes slik at datamaterialet er anonymisert.

Dine rettigheter.
Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke personopplysninger som er registrert om deg, og å få utlevert en kopi av opplysningene,
- å få rettet personopplysninger om deg,
- å få slettet personopplysninger om deg, og
- å sende klage til Datatilsynet om behandlingen av dine personopplysninger.

Hva gir oss rett til å behandle personopplysninger om deg?
Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra Institutt for språk og litteratur ved NTNU har NSD (norsk senter for forskningsdata) vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Hvor kan jeg finne ut mer?
Hvis du har spørsmål til studiet, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- Institutt for språk of litteratur ved Erlend Rannem erlendnr@ntnu.no og/eller Mila Vulchanova mila.vulchanova@ntnu.no.

- Vårt personvernombud: Thomas Helgesen thomas.helgesen@ntnu.no.

Hvis du har spørsmål knyttet til NSD sin vurdering av prosjektet, kan du ta kontakt med:

- NSD – Norsk senter for forskningsdata AS på epost (personverntjenester@nsd.no) eller på telefon: 55 58 21 17.

Med vennlig hilsen

Mila Vulchanova	Erlend Ness Rannem
(veileder)	(student)

Appendix 4: Consent form

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet *«Andrespråksbrukeres prosessering av engelske setninger»*, og har fått anledning til å stille spørsmål. Jeg samtykker til:

- ☐ å delta i forskningsprosjektet

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet.

(Signert av prosjektdeltaker, dato)

Appendix 5: Background questions

Del A: Personlig informasjon

Studieprogram: ____

Fødselsår: __________________

Kjønn: □ Mann □ Kvinne

Del B: Språklig bakgrunn

I engelsk, hvordan vurderer du ferdighetene dine i hvert av disse områdene?

	Grunnleggende	Middels	Avansert	Flytende
Lesning				
Skriving				
Snakke				
Lytte				
Totalt				

Når begynte du å tilegne deg engelsk? (f.eks. årstall eller klassetrinn) __________________

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/hører du engelsk?

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

Appendix 6: The English sentence completion test

Complete the sentences

In this part, you will be presented with various incomplete English sentences. Your task is to form complete sentences by filling in the blanks.

The only rule is that your answers must begin with *THE, A,* or *AN*, followed by a noun (substantiv).
See the examples below.

Example 1

Ben drove [blank] all the time.

the car

Example 2

The woman found [blank] during a hiking trip.

an apple

Example 3

He quit [blank] last semester.

the course

Example 4

Willis arrested [blank] in 1989.

a man

No.	Lemma	Sentences			
		Noun	Verb	[blank] *your answer*	Adverbial
1	Attend	Mitch	attended		daily.
2	Board	Emma	boarded		the fastest.
3	Boil	Roy	boiled		for far too long.
4	Broadcast	Jerome	broadcasted		all day.
5	Button	Scotty	buttoned		rapidly.
6	Buy	Kevin	bought		last Monday.
7	Carry	He	carried		practically.
8	Cherish	The woman	cherished		truthfully.
9	Cite	Harold	cited		far too often.
10	Criticize	Travis	criticized		every day.
11	Decorate	Shelly	decorated		every other month.
12	Dial	Felix	dialed		as soon as he woke up.
13	Dig	Andrew	dug		quite loudly.
14	Drink	Ollie	drank		in haste.
15	Dump	Carl	dumped		two days ago.
16	Eat	Ralph	ate		in an awkward manner.
17	Elect	He	elected		carefully.
18	Entertain	Bert	entertained		to a certain degree.
19	Extinguish	Edgar	extinguished		in 1999.
20	Fire	Lincoln	fired		cheerfully.
21	Flush	David	flushed		hastily.
22	Fry	Sid	fried		to perfection.
23	Guard	I	guarded		for six days straight.
24	Hammer	I	hammered		without care.
25	Join	Nathan	joined		a long time ago.
26	Kill	He	killed		in 2019.
27	Know	She	knew		finally.
28	Land	Moe	landed		just now.
29	Light	Nevil	lighted		with a smile on his face.
30	Lock	Ross	locked		as one should.
31	Marry	The man	married		moments later.
32	Memorize	Quinn	memorized		without trying.
33	Mend	Maxwell	mended		to the best of his effort.
34	Milk	Lily	milked		at sunrise.
35	Move	Sylvester	moved		a minute ago.
36	Mow	Curtis	mowed		outside his house.
37	Order	Billie	ordered		by impulse.

		Sentences			
No.	Lemma	Noun	Verb	[blank] *your answer*	Adverbial
38	Plant	The woman	planted		far away.
39	Play	She	played		more often than not.
40	Predict	Leo	predicted		by accident.
41	Prune	Karen	pruned		thoroughly.
42	Publish	Zack	published		last April.
43	Push	Alf	pushed		lazily.
44	Read	I	read		in the living room.
45	Rent	He	rented		from time to time.
46	Rub	Lucas	rubbed		at 4 o'clock.
47	Saddle	Robert	saddled		a minute ago.
48	Sail	Harry	sailed		as if he was an expert.
49	Smoke	Edward	smoked		with delight.
50	Solve	She	solved		yesterday.
51	Sow	Irwin	sowed		three days ago.
52	Tap	The man	tapped		hardly.
53	Throw	He	threw		over there.
54	Tie	Eddie	tied		in a hurried manner.
55	Want	Eric	wanted		more than anything else.
56	Wash	She	washed		until she got bored.
57	Watch	Timmy	watched		whilst in New York.
58	Water	Stacy	watered		a day later than expected.
59	Wear	She	wore		elegantly.
60	Write	She	wrote		in France.

Appendix 7: LexTALE

Vocabulary test

You will now be presented with several "words", one at a time. Your task is to decide whether this word is an existing English word or not. If you think it is an existing English word, you click on *YES*, and if you do not think it is an existing English word, you click on *NO*.

If you are sure that the word exists, even though you do not know its exact meaning, you may still click on *YES*. But if you are not sure if it is an existing word, you click on *NO*.

The words are written in British English rather than American English spelling. For example: "realise" > "realize"; "colour" > "color", and so on. Do not let this confuse you. This experiment is not about detecting such subtle spelling differences anyway.

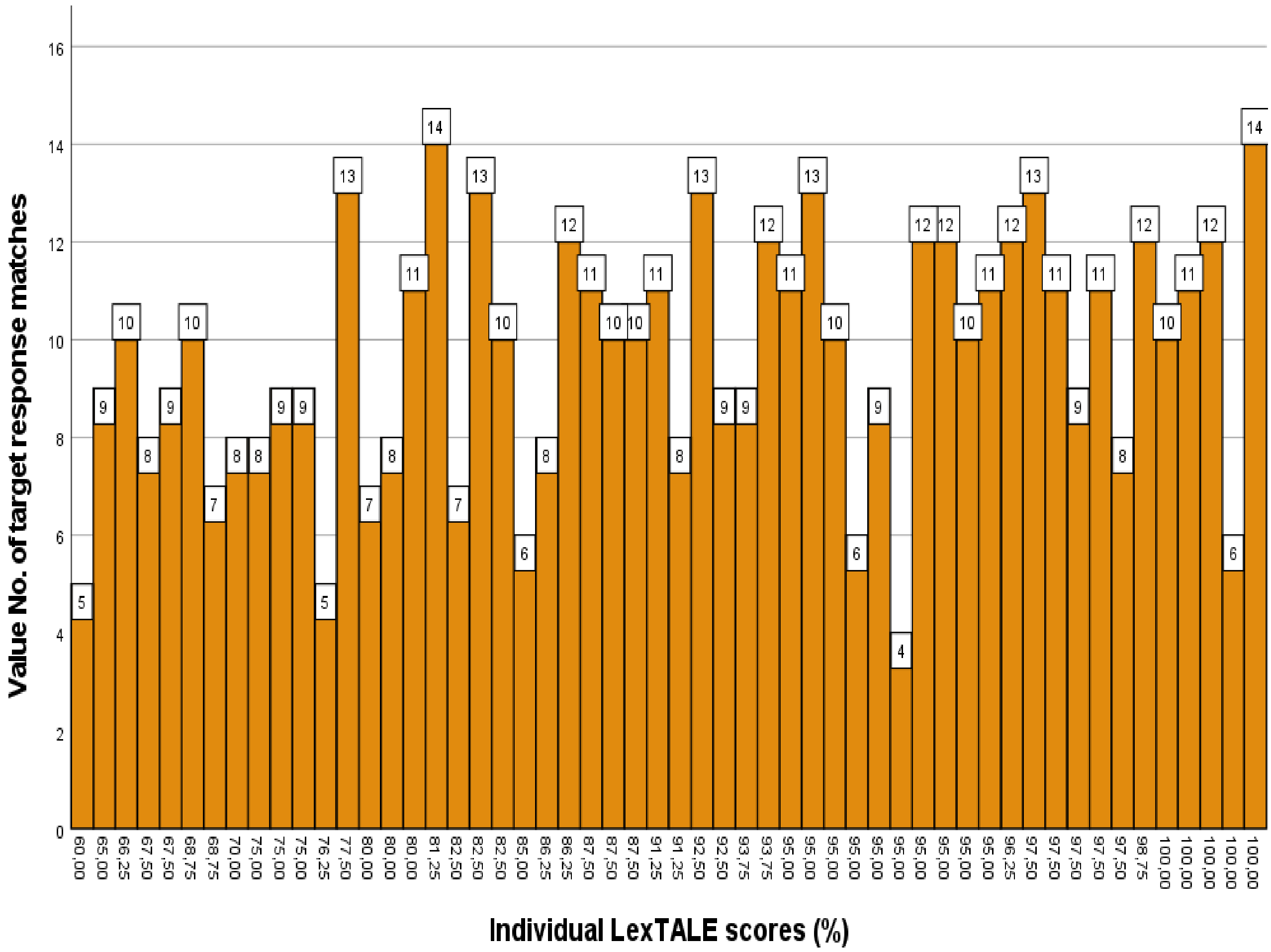
No.[26]	Item	*your answer*		Word status[27]
0	*platery*	□ YES	□ NO	0
0	*denial*	□ YES	□ NO	1
0	*generic*	□ YES	□ NO	1
1	*mensible*	□ YES	□ NO	0
2	*scornful*	□ YES	□ NO	1
3	*stoutly*	□ YES	□ NO	1
4	*ablaze*	□ YES	□ NO	1
5	*kermshaw*	□ YES	□ NO	0
6	*moonlit*	□ YES	□ NO	1
7	*lofty*	□ YES	□ NO	1
8	*hurricane*	□ YES	□ NO	1
9	*flaw*	□ YES	□ NO	1
10	*alberation*	□ YES	□ NO	0
11	*unkempt*	□ YES	□ NO	1
12	*breeding*	□ YES	□ NO	1
13	*festivity*	□ YES	□ NO	1
14	*screech*	□ YES	□ NO	1
15	*savoury*	□ YES	□ NO	1
16	*plaudate*	□ YES	□ NO	0

[26] Item number. (Note that the first three items are dummies.)
[27] Word status; 0 = nonword, 1 = word

No.	Item	your answer		Word status
17	*shin*	□ YES	□ NO	1
18	*fluid*	□ YES	□ NO	1
19	*spaunch*	□ YES	□ NO	0
20	*allied*	□ YES	□ NO	1
21	*slain*	□ YES	□ NO	1
22	*recipient*	□ YES	□ NO	1
23	*exprate*	□ YES	□ NO	0
24	*eloquence*	□ YES	□ NO	1
25	*cleanliness*	□ YES	□ NO	1
26	*dispatch*	□ YES	□ NO	1
27	*rebondicate*	□ YES	□ NO	0
28	*ingenious*	□ YES	□ NO	1
29	*bewitch*	□ YES	□ NO	1
30	*skave*	□ YES	□ NO	0
31	*plaintively*	□ YES	□ NO	1
32	*kilp*	□ YES	□ NO	0
33	*interfate*	□ YES	□ NO	0
34	*hasty*	□ YES	□ NO	1
35	*lengthy*	□ YES	□ NO	1
36	*fray*	□ YES	□ NO	1
37	*crumper*	□ YES	□ NO	0
38	*upkeep*	□ YES	□ NO	1
39	*majestic*	□ YES	□ NO	1
40	*magrity*	□ YES	□ NO	0
41	*nourishment*	□ YES	□ NO	1
42	*abergy*	□ YES	□ NO	0
43	*proom*	□ YES	□ NO	0
44	*turmoil*	□ YES	□ NO	1
45	*carbohydrate*	□ YES	□ NO	1
46	*scholar*	□ YES	□ NO	1
47	*turtle*	□ YES	□ NO	1
48	*fellick*	□ YES	□ NO	0
49	*destription*	□ YES	□ NO	0
50	*cylinder*	□ YES	□ NO	1
51	*censorship*	□ YES	□ NO	1
52	*celestial*	□ YES	□ NO	1
53	*rascal*	□ YES	□ NO	1
54	*purrage*	□ YES	□ NO	0
55	*pulsh*	□ YES	□ NO	0

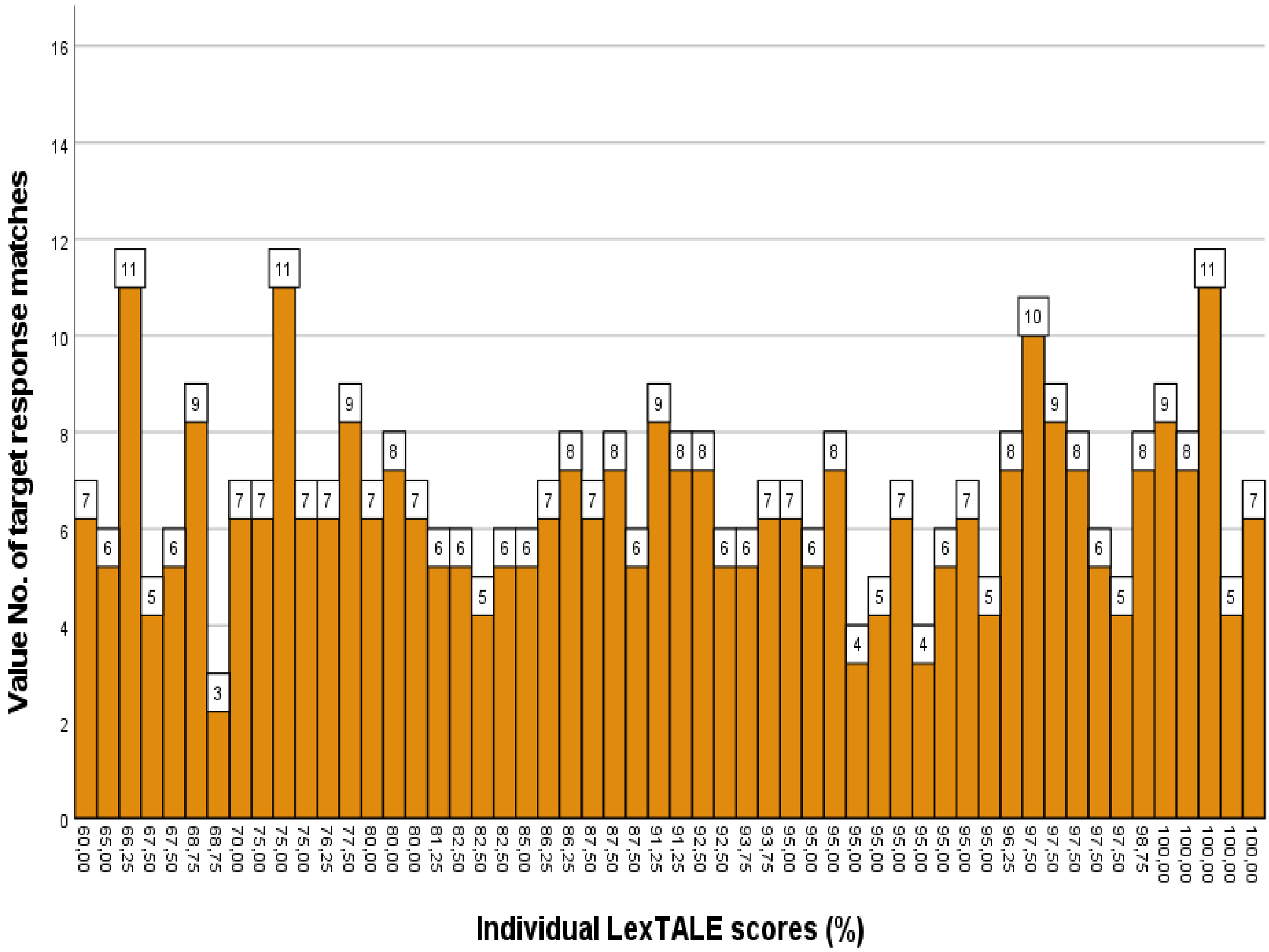
No.	Item	*your answer*		Word status
56	*muddy*	□ YES	□ NO	1
57	*ouirty*	□ YES	□ NO	0
58	*pudour*	□ YES	□ NO	0
59	*listless*	□ YES	□ NO	1
60	*wrought*	□ YES	□ NO	1

Appendix 8: Individual target response matches in the most constrained category



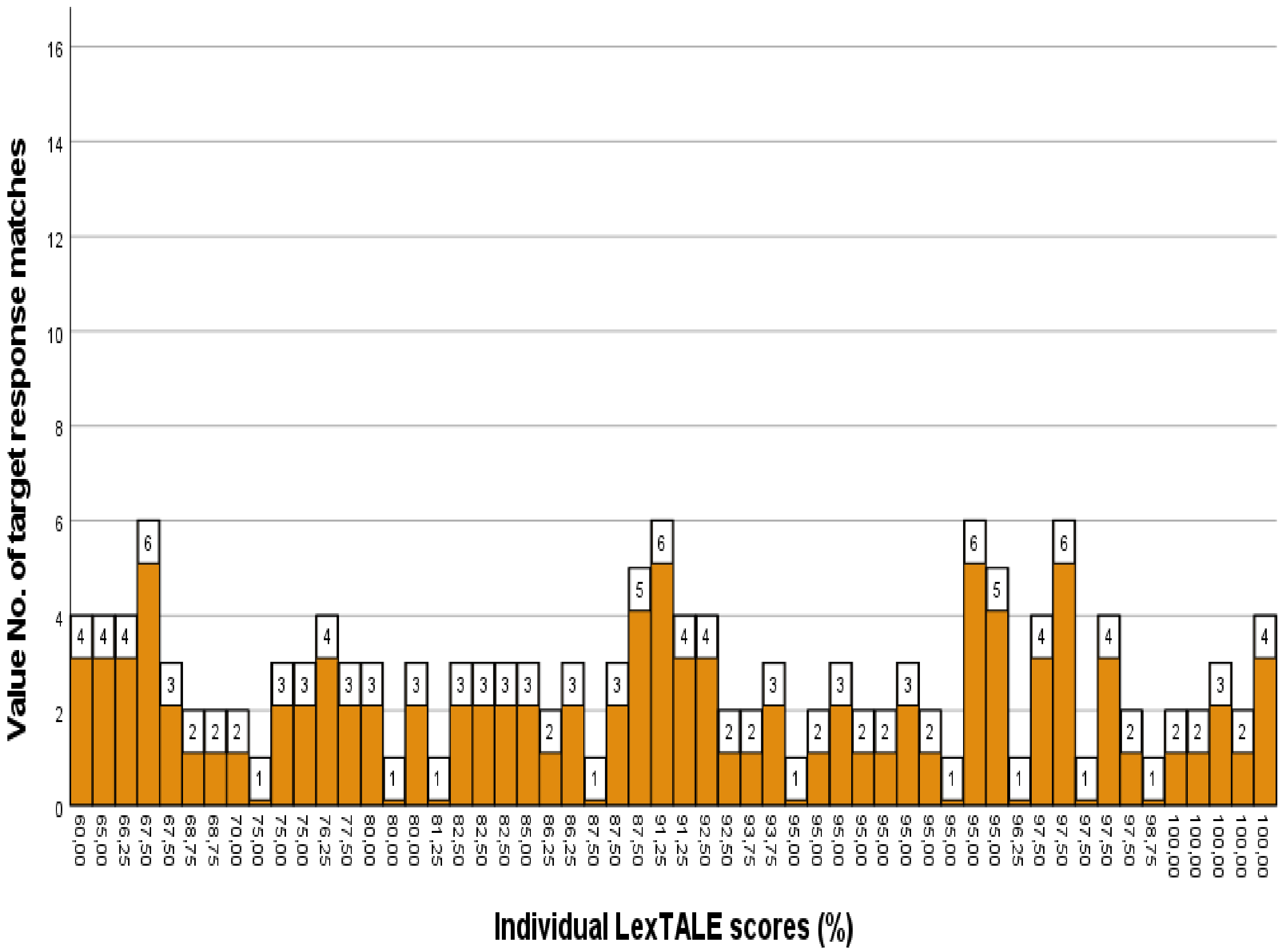
Note: The figure encompasses both the frequent and infrequent verbs of the most constrained category. As a result, scores range from 0 to 20.

Appendix 9: Individual target response matches in the moderately constrained category



Note: The figure encompasses both the frequent and infrequent verbs of the moderately constrained category. As a result, scores range from 0 to 20.

Appendix 10: Individual target response matches in the least constrained category



Note: The figure encompasses both the frequent and infrequent verbs of the least constrained category. As a result, scores range from 0 to 20.

Appendix 11: Unique responses for verbs addressed in Section 7.2

Boil	Number of responses	Land	Number of responses	Elect	Number of responses
egg	33	plane	29	president	18
water	4	deal	7	candidate	4
pasta	3	job	4	man	3
rice	3	shot	3	option	2
potato	2	trick	2	person	2
spaghetti	2	balloon	1	politician	2
apple	1	drone	1	administrator	1
chicken	1	gig	1	clothes	1
drink	1	helicopter	1	council	1
frog	1	jump	1	coworker	1
meat	1	part	1	elect	1
stew	1	position	1	gift	1
yam	1	spaceship	1	girl	1
		spell	1	judge	1
				kiwi	1
				office	1
				one	1
				pencil	1
				piece	1
				puppy	1
				representative	1
				seat	1
				sofa	1
				strategy	1
				supervisor	1
				team	1
				winner	1
				woman	1
				word	1

Appendix 12: The Norwegian sentence completion test

Fullfør setningene

Du vil nå bli presentert med ufullstendige setninger.
Din oppgave er å fullføre disse ved å fylle inn tomrommet i slutten av hver setning.

Den eneste regelen er at hvert svar må

- **ENTEN** være i bestemt form (se Eksempel 1)
- **ELLER** være i ubestemt form entall (se Eksempel 2)

I tillegg, prøv å ikke overtenke svaret, men skriv heller det første du kommet på.

Eksempel 1

Karin kastet …

eplet / eplene

Eksempel 2

Johann knuste …

en / ei / et vindu

FORTSETT

Eirik kokte …

[ ]

Appendix 13: Unique responses for verbs addressed in Section 7.3

Sow	Number of responses	Mow	Number of responses	Prune	Number of responses
seed	24	lawn	37	bush	14
blanket	3	grass	6	tree	9
dress	3	chair	3	*blank*	3
plant	2	bin	1	flower	3
sweater	2	blanket	1	hedge	3
wheat	2	cat	1	branch	2
bean	1	couch	1	grape	2
field	1	dog	1	plant	2
flower	1	flower	1	prune	2
ground	1	plant	1	database	1
handkerchief	1	truck	1	dog	1
jeans	1			friendship	1
key	1			horse	1
lawn	1			log	1
picture	1			meat	1
piece	1			paper	1
potato	1			pearl	1
riddle	1			rose	1
seam	1			text	1
shirt	1			thing	1
skirt	1			thought	1
tomato	1			tulip	1
tree	1			vegetable	1
wound	1				

NTNU

Norwegian University of
Science and Technology