Master's thesis

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

Stefan Williams

U-shaped Trajectories in L2 Learning: Testing the Dual Processing Hypothesis

An Experimental Study of Norwegian L2 Learners of English

Master's thesis in Language Studies with Teacher Education - Master's Program in English
Supervisor: Mila Dimitrova Vulchanova
December 2019

NTNU
Norwegian University of Science and Technology

Stefan Williams

U-shaped Trajectories in L2 Learning: Testing the Dual Processing Hypothesis

An Experimental Study of Norwegian L2 Learners of English

Master's thesis in Language Studies with Teacher Education - Master's Program in English
Supervisor: Mila Dimitrova Vulchanova
December 2019

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

NTNU
Norwegian University of
Science and Technology

Abstract

This study explores U-shaped behaviour in the acquisition of irregular verb morphology across three different groups of Norwegian L2 learners of English. This phenomenon is of special interest due to its significance for the organization and division between the mental lexicon and grammar. A cross-sectional design with randomly recruited students from three different grades was employed. The final analysis includes data from participants within the 8th grade (N=17), 9th grade (N=19), and 10th grade (N=15). We report results on the acquisition of irregular verb and noun morphology, in addition to mean reaction times on different types of responses (accurate responses and overregularized ones). The students were given elicitation tasks based on Berko's (1958) Wug-test to mainly test irregular past tense verbs via the use of an online survey-platform. We hypothesized that if a U-shaped trajectory could be observed across the three groups, then we would find significant differences between the three groups in accuracy levels in conjunction with overregularization errors. Our results model the later stages of U-shaped learning where we found an increase in overall accuracy co-occurring with a decrease in overregularization errors. We propose that the existence of U-shaped behaviour in the L2 suggests universality as opposed to simply L1 transfer, and that this learning process is an integral part of acquiring and establishing knowledge in the mental lexicon related to regular and irregular morphology. Finally, we theorize how U-shaped behaviour speaks in favour of a dual-organization of the mental faculty as opposed to a connectionist account.

Acknowledgements

First, I would like to thank my supervisor Professor Mila Dimitrova Vulchanova. Your guidance has been invaluable as well as inspiring while embarking on this process of writing my thesis. Moreover, thank you for giving me the opportunities that you have during my time as a student to learn and develop as a researcher. I would also like to thank Professor Giosuè Baggio for helping me with the statistics for my thesis. This thesis would not be the same without your help.

I would like to thank my friends and family for showing an interest and supporting me during the past one and a half year of writing and working on this project. I am particularly grateful of my fiancée for being extra supportive during the writing process, and of my father for motivating and inspiring me despite his illness. I could not have done it without you.

Special thanks to Olav Landmark Pedersen for helping me design the plots and figures in R, and to Ellen Saxlund and Iris Kummen for giving me insightful feedback throughout the course of writing the thesis. Your comments and encouragement have really made a difference. Lastly, thank you to all the teachers and students involved in the data-collection of this study, and to the baseball community in Trondheim for giving me an environment to relax and get a break from work.

Contents

Figures

Tables

Abbreviations/symbols

ASD	Autism spectrum disorder
ERP	Event-related potentials
L1	First language
L2	Second language
LAN	Left anterior negativity
NSD	Norwegian Centre for Research Data
SLA	Second language acquisition
VG3	Third year of upper secondary school in Norway

1 Introduction

Phenomena such as grammatical errors and overregularization have been of interest for researchers within the field of language acquisition for a long time, especially in relation to inflectional morphology of verbs. As pointed out by Marcus et al. (1992, p. 1), overregularization of irregular verbs, such as *comed* instead of *came* for instance, has been commented on for as long as language development has been a topic of study, going as far back as Bateman (1916).

Closer studies of this particular type of error revealed an interesting developmental curve. It was first reported by Ervin & Miller (1963; see also Cazden 1968) and is currently referred to as the *U-shaped developmental curve*. Children with English as their L1 produce correct irregular forms of verbs, if they mark the past tense at all, up until their third year. However, around the age of three, they start producing errors and overregularize the irregular verbs that they previously have been able to mark correctly. Around the same time as they start producing these errors, children begin acquiring the rule for marking regular past tense verbs (Pinker, 1998).

To the best of my knowledge, research within language acquisition on the topic of the U-shaped developmental curve has mainly focused on L1 acquisition. This study explores whether or not there is any evidence of U-shaped learning in L2 acquisition in the same domain it has been attested in L1-tense morphology of verbs. If so, I question if it there is any evidence for a dual-organization, as proposed by Pinker (1991, 1998), of the mental lexicon and grammar of L2 speakers of English.

To do this, elicitation tasks based on Berko's (1958) Wug-test were designed to elicit the participants' knowledge of irregular morphology of both verbs and nouns. Data were collected via the online survey platform *SurveyGizmo*. Norwegian speakers with English as their L2 were chosen as the focus group for this study. Given the time constraints of this MA-project, a cross-sectional design was chosen, recruiting students from three different grades, respectively the 8th grade (N=20), 9th grade (N=24), and 10th grade (N=23)[1].

We expected to find significant differences in accuracy and overregularization rates between the three groups. If we do indeed find evidence of U-shaped learning, it would speak in favour of the universality of U-shaped learning as opposed to simply L1 transfer. We hypothesize that if L2 speakers are indeed influenced by positive transfer from their L1, Norwegian, which displays the same phenomenon (strong verb classes), then they would be simply storing the irregular forms of verbs in the L2 (English) in long term memory, thus exhibiting a steady increase in accuracy. There would be no patterns of overregularization nor any U-shaped behaviour. Further, evidence for U-shaped learning would also suggest a dual-processing mechanism which not only underlies L1 acquisition, but also L2 acquisition. The possibility of finding ceiling effects was also taken into consideration.

[1] Number of participants prior to the exclusion of outliers and other non-target participants.

2 Previous Research

2.1 The U-shaped developmental curve: An introduction based on findings from previous research

In introducing the theoretical background for the current thesis, I begin by addressing the U-shaped developmental curve, also referred to as *U-shaped learning* or *U-shaped acquisition*. The existing literature on U-shaped learning with respect to SLA appears to be scarce, which is why most of the literature referred to in this chapter pertains to first- rather than second language acquisition.

Overgeneralization errors made by children linked to irregular verbs, such as producing *felled* instead of *fell*, have been observed and contemplated upon as far back as Bateman (1916). In describing the language produced by three children (L1), Bateman (1916) observed how two of the three children "[…] did use many past tense forms, especially Anne, but they had considerable difficulty with irregular forms so that 'seed,' 'felled,' 'falled,' runned,' etc., were common in their speech" (Bateman, 1916, p. 229).

Later studies on overgeneralization errors of irregular verbs in L1 production revealed a universal developmental trajectory. In a small longitudinal study, Cazden (1968) observed similar errors as those reported by Bateman (1916). Furthermore, the children exhibited a period of correct performance when producing the irregular past tense verbs prior to the overgeneralization errors. The same phenomenon was observed by Cazden (1986) in relation to irregular nouns where the participants applied the productive rule to an irregular stem resulting in forms such as *feets* instead of *feet*. Moreover, there seemed to be individual differences in the error rate among participants. The same pattern of performance was later supported by a larger study utilizing data from spontaneous language produced by 83 children (Marcus et al., 1992). The study gave further insight into the acquisition of rule-like behaviour relating to past tense morphology, in addition to the affirmation of the U-shaped learning curve.

Marcus' and colleagues' (1992) study of overgeneralization errors established seven findings, which are as follows. (1) Overregularization of verbs occurs infrequently, about median 2.5% of irregular verbs, which in turn suggests that there is no defect in the child's grammar that must be unlearned. (2) These errors occur at a low rate between age 2 and into early school years, affecting most irregular verbs. (3) Children exhibit an extended period of correct performance before the first error. (4) There is no correlation between overregularization and the increase in number of regular verbs in the parental speech, which counteracts the connectionist account where statistical sensitivity plays an essential role in adopting rules. (5) Overregularization first occurs when the child begins to mark the regular verbs for tense consistently. (6) The more a child is exposed to the irregular form of a verb from their parents, the less the child tends to overregularize the given verb. (7) Similar-sounding irregular verbs are not affected by overregularization, but they are also not attracted to overregularization by similar-sounding regulars. This suggests that irregular patterns are stored in associative memory with connectionist

properties, whereas regulars are not (Marcus et al., 1992). Based partly on these findings and previous work by Pinker and colleagues (Pinker, 1991; Pinker & Prince, 1988), the following hypothesis about overregularization errors seems reasonable. Children exhibit a period of correct performance when producing irregular verbs, if they mark the past tense at all. Lasting from the age of two and into early school years, errors are produced in the form of overgeneralizing the regular inflection of the past tense, resulting in the overregularization of irregular verbs that were previously correctly inflected. However, with time, they reverse back and start marking the irregular verbs correctly again. This dip in performance, which is where the U-shaped developmental curve gets its name from, seems to affect most irregular verbs. The phenomenon co-occurs with the acquisition of the rule of regular past tense inflection of verbs (Pinker, 1991, 1998; Pinker & Price, 1988; Marcus et al., 1992).

An important point of clarification made by Marcus et al. (1992, p. 40-44) was defining what the phenomenon U-shaped learning refers to. Their study relates to Cazden's (1968) and Rumelhart & McClelland's (1986) definition where they conceptualize the phenomenon as "[…] a transition from a period in which past tense forms are marked correctly whenever they are marked at all to a period in which some overregularization errors occur as well" (Marcus et al., 1992, p. 40). This is the same definition as the one described in the previous paragraph, and the one which I will relate to. There are other senses of the U-shaped curve which refer to other phenomena. As pointed out by Marcus et al. (1992):

> In Plunkett & Marchman (1991), the learning curves all start out at levels of performance far less than 100% and then increase; the authors call the small wiggles in this overall increasing curve "U-shaped development." Although all the children we examine show local ups and downs in their monthly measures of overregularization rates, there are many explanations of these blips, of which sampling errors is the simplest. (Marcus et al., 1992, p. 44)

As the quote illustrates, there are other senses of the U-shaped curve which refer to different phenomena. As a matter of delineation, these short-term fluctuations will not be considered as U-shaped learning in the sense discussed in this thesis.
U-shaped learning is not only limited to language acquisition. From the cognitive-developmental literature, Carlucci & Case (2013) report that U-shaped learning occurs in a variety of child-developmental phenomena such as understanding temperature, weight conservation, object permanence, and facial recognition. In questioning whether U-shaped learning is logically necessary or not in relation to some formal learning tasks, Carlucci & Case (2013) contend in their paper that:

> […]the general picture that emerges from the so-far known results presented in this article is that U-shaped behavior is unavoidable for full learning power in the context of a number of parametrized models of learning featuring a number of cognitively motived constraints. The results might be taken as suggestive of the fact that humans might exhibit U-shaped and other nonmonotonic learning patterns[…]. (Carlucci & Case, 2013, p. 58)

Marcovitch & Lewkowicz (2004) pose the question if U-shaped developmental trajectories are more than an interesting artefact of developmental processes. Rather, they ask if these curves are in fact a hallmark of such processes. Drawing parallels between language acquisition and ontological adaptions as proposed by Openheim (1981, as cited in Marcovitch & Lewkowicz, 2004, p. 115), they postulate how this kind of behavioural

regression is central for the developmental process itself. Given the aforementioned articles, it is an interesting question whether or not U-shaped learning might indeed be a universal pattern in human cognitive development.

In sum, the U-shaped pattern involves the following. There is a decline in accuracy for the production of irregular verbs which have previously been marked correctly. Co-occurring with this drop in accuracy are an increase in overregularization errors of irregular verbs. With time and as development continues, the rate of overregularization errors decreases and the individual goes back to marking the irregular verbs consistently again. Given this, it is conceivable that when observing U-shaped learning, depending on the interval length of time that one observes it, one might only observe sections on the phenomenon. For instance, a drop in accuracy coinciding with an increase in overregularization errors would suggest earlier stages of U-shaped learning. Conversely, a decrease in overregularization errors taking place simultaneously with an increase in overall accuracy would be indicative of a later stage of U-shaped learning. Alternatively, only an increase in overregularization errors might suggest that one is observing the mid-section.

2.2 Theoretical perspectives on U-shaped learning

U-shaped learning as a phenomenon has been a point of discussion on the controversial topic of the psychological reality of the organization of the brain, how the mental lexicon is organized and how it operates (cf. Rumelhart & McClelland, 1986; Plunkett & Marchman, 1991; Pinker & Prince, 1988; Marcus et al., 1992; Pinker 1998). On the one hand, there is the traditional account within linguistics which claims that language is the product of an interaction between memorized entities and symbol manipulation. Phrased differently, it is the interaction between lexical items and the underlying rules that accounts for the expressive power of language. This involves the extensive integration of typed variables such as the phoneme /b/ or N as a category for instance. On the other hand, there is the connectionist account which seeks to make memory more powerful and to model language based on a neurological perspective. The theory claims that there are in fact no rules and that language is the result of pattern associations and spreading activation throughout networks. As an alternative to both these two theories, combining aspects of both, the dual-mechanism theory maintains that there are two modules which govern language, just as the traditional account holds. However, the mental lexicon, as a form of memory, has associative properties as proposed by the connectionist account (Pinker, 1991). In the following section, an overview of the theories, accounts and discussions surrounding U-shaped development, with a special emphasis on the dual-mechanism theory, will be presented.

The traditional account of overgeneralizations relies on the disassociation between the two following psychological processes; rote memory and the deployment of rules. In the course of developing language, children primarily memorize the different verb forms that they hear. If they hear *spoke*, they consequently produce *spoke*. Over time as they receive more input, they form the regular past tense rule by abstracting the pattern from the available data accumulated over time. The younger children, prior to overregularizing irregulars, do not apply the regular past tense suffix since they have not yet acquired it. Older children, however, do apply the productive rule to irregulars since they have in fact

acquired the rule, which consequently lead to overregularization errors (Marcus et al., 1992, p. 6).

Authors such as Pinker and colleagues (Pinker, 1991, 1998; Marcus et al., 1992) contend that the traditional account is inadequate since it is incapable of accounting for certain behavioural observations and patterns among irregular verbs. There are sub-regular patterns among irregular verbs such as *sing-sang, ring-rang, shrink-shrank*, which are problematic to account for within this theoretical framework. Pinker (1991) contends that "The rote memory cannot explain why verbs with irregular past forms come in families, rather than belonging to arbitrary lists" (Pinker, 1991, p. 531). Moreover, adults seem capable of inflecting nonsense verbs into irregular sounding past tense forms, such as the pair *spling-splung*, due to the phonological conditions of the present tense form (Bybee & Moder, 1983).

Pinker (1998) reports Bybee's (1985, as cited in Pinker, 1998, p. 11) finding that Old English used to contain twice as many irregular verbs as today's Modern English, meaning that there used to be irregular (strong) verbs that are now obsolete such as *cleave-clove*. Today's irregular verbs used to belong to older productive paradigms which were based on their (morpho)phonological properties. These older productive inflectional paradigms have disappeared over time due to the process of attrition and simplification. Although today's irregular verbs used to belong to these previously productive paradigms, they must be stored in memory as separate items. However, one could argue that people's ability to form irregular sounding novel words such as *spling-splung* is the result of tapping into the (morpho)phonological structure of the word, and thus also in a sense tapping into the older productive inflectional systems that are now obsolete in Modern English.

There are several studies which support this idea besides the decline in the number of irregular verbs that we see over time. For instance, Boudelaa & Marslen-Wilson (2011, as cited in Vulchanova, Talcott, Vulchanov, Stankova & Eshius, 2013, p. 587) demonstrated that there are abstract representations of roots which facilitate pattern extraction. Vulchanova et al. (2013) report how there are recent studies which suggest "[…] that word processing is determined by its (morpho)phonological structure and whether this structure provides the language user with a structural pattern rather than the regularity status of the word" (Vulchanova et al., 2013, p. 587). Furthermore, Vulchanova et al. (2013) did a case study of an L1 speaker of Bulgarian with ASD (Asperger's), comparing among other things the participant's knowledge of morphology against a neurotypical control group (N=20). They found differences between the I.A. (participant with ASD) and the control group showing how the I.A. outperformed the control group, especially in relation to irregular verbs. In interpreting these results, they hypothesized that the I.A. processes regular and irregular words in a similar way relying on pattern extractions permitted by phonologically conditioned grammatical paradigms in Bulgarian. The control group, on the other hand, may be utilizing a different system (Vulchanova et al., 2013, p. 595). The results could indicate that the newer generations of Bulgarian L1 speakers are in fact not acquiring the sub-regular patterns of irregular verbs, thus indicating a similar diachronic language development as with Old English to Modern English. Furthermore, it is conceivable that this other system, utilized by the control group, is analogous to a dual-mechanism system.

In challenging the traditional conceptualization of the organization of the language faculty, Rumelhart & McClelland (1986, as cited in Ellis, 2003) proposed a connectionist approach to language acquisition based on a neural inspired learning model. These connectionist learning models are today referred to as sub-symbolic systems in artificial intelligence-theory, and do not operate on symbols, but rather only on connections and networks. Interestingly, Rumelhart & McClelland's (1986) model managed to generate the U-shaped learning curve in acquiring the English past tense forms. Connectionist theory (Ellis, 2003) hypothesizes that language is the result of connections between processing units which in turn form larger networks that are interconnected into one large unit. Within this theoretical framework, language operates within one interconnected module, as opposed to the traditional account which posits the view that there are two modules (the mental lexicon and the grammar module) which function simultaneously to form language. In network-based language models, there are no rules which govern language. Rather, only associations do so which in turn means that language is in fact governed by connections between nodes and the networks alone. The theory considers the learning process in both L1 and L2 acquisition as being the same. From the theoretical perspective of connectionism, as Ellis (2003, p. 87) has pointed out, the question is to what degree these two processes are similar and which limits are involved are unknown.

Connectionist learning models, such as the aforementioned model by Rumelhart & McClelland (1986, as cited in Ellis, 2003), rely on the frequency of language chunks in the input data to acquire language, thereby adjusting the weights of the nodes. One of the advantages of connectionist models and theory is how they are "[…] data rich and process-light. Massively parallel systems of artificial neurons use simple learning processes to statistically abstract information from masses of input data" (Ellis, 2003, p. 85). These models seem capable of capturing regularities and patterns found in language. This means that they can extract the regularities and consequently function in a rule-like way, accounting for the descriptive regularities found in language. However, this is not the same as considering language as being governed by rules, as conceived by the traditional account. There is also the matter of using variables in linguistic theories in relation to the *binding problem*[2], a problem commented on by authors such as Jackendoff (2009, p. 64), which does pose a challenge for cognitive neuroscience (Jackendoff, 2009, p. 58). Rumelhart & McClelland's (1986) connectionist model circumvents this matter entirely in the way that the model is designed, and it is still capable of reproducing, although arguably not perfectly, a U-shaped developmental curve. Conversely, not incorporating variables is also one of the criticisms towards network models in the larger discussion between psycholinguists and neurolinguists on the topic of the psychological reality of the organization of the language faculty (Jackendoff, 2009). I will briefly readdress the matter of variables in linguistic theory later in this chapter.

Although there are positive aspects of connectionism and connectionist models, the psychological reality of them have been challenged extensively by authors such as Pinker

[2] Jackendoff (2002, p. 59) writes: "The need for combining independent bits into a single coherent percept has been recognized in the theory of vision under the name of the *binding problem*[…]". In relation to linguistics, it relates to the problem of conceptualizing and modelling language in such a way that one can capture the interface between all the different relations in language (semantic, syntactic, phonological) in a way that consolidates our theories of language with our knowledge and understanding of neural networks. Adding a temporal aspect to such models is still a challenge for our understanding of language.

& Prince (1988) and Marcus et al. (1992). For instance, the Rumelhart & McClelland (1986) model can learn mappings found in no human language, and at the same time it has difficulties learning common mappings such as reduplicating stems (Pinker 1991; Pinker & Prince, 1988):

> Lacking in representation of words as lexical entries, distinct from their phonological or semantic content, the model cannot explain how languages can contain semantically unrelated homophones with different past tense forms such as *lie-lied* (prevaricate) and *lie-lay* (recline), *ring-rang* and *wring-wrung*, *meet-met* and *mete-meted.* (Pinker, 1991, p. 531)

The Rumelhart & McClelland (1986) learning model also failed at producing output for many novel regular verbs that were dissimilar to the ones in the training set (Pinker, 1991). Considering how readily people can produce regular sounding novel verbs, this does pose a problem to the learning model. These are three examples that illustrate how there are problematic elements to the Rumelhart & McClelland-model. For further criticism, see Pinker & Prince (1988), and Marcus et al. (1992). There have been revised models based on Rumelhart & McClelland's (1986) learning model, such as Plunkett & Marchman's (1991) connectionist model. As mentioned earlier, instead of creating a U-shaped learning curve, as observed and discussed in the literature on U-shaped learning (Cazden, 1968; Pinker, 1991; Marcus et al., 1992; Pinker 1998; Karmiloff & Karmiloff-Smith, 2002), Plunkett & Marchman's (1991) model created short-term fluctuations in a steadily increasing curve of performance which they refer to as U-shaped learning (Marcus et al., 1991, p. 44). It is thus conceivable that connectionist models need further improvement if connectionism is to be considered a viable alternative to more traditional theories of the composition of the language faculty. This seems to be the case at least for the ones that are similar to the aforementioned models from the 80s and 90s. Jackendoff (2009, p. 165) proposes that network models should accept the challenge of integrating typed variables into their models:

> […] if a multi-layer network trained by back-propagation in principle cannot account for something as combinatorially trivial as the regular past tense, then there is no hope of scaling up current connectionist solutions to the past tense of the rest of the language, where free combination reigns supreme. Rather, […], the challenge to network approaches is to develop a robust device for encoding variables – not to continue to try to live without them. (Jackendoff, 2009, p. 165)

Although he makes this critical remark towards connectionist models, Jackendoff (2002) also hypothesizes how the dual-mechanism is problematic for incorporating variables. However, due to the scope of this thesis, I will not pursue this any further. He concludes that finding a way of integrating variables into network-based models would lead to a very fruitful dialogue between the network modelers and the theoretical linguists.

Connectionist models have especially been subject to criticism by proponents of the dual-mechanism theory. Dual-mechanism, as described by Pinker (1991, 1998), combines aspects of the traditional account of the interaction between memory and rules, with associative properties from connectionist theory in regard to memory and the mental lexicon. It seeks to out to create a theory which has the ability to account for the shortcomings of the traditional account and connectionist models.

Similar to the traditional conceptualization, the dual-mechanism theory proposes that there are two modules which interact with each other to form language. There is the memory component which comprises the mental lexicon, storing words in the form of either stems or the full irregular form. The second module is the rule-capturing component which governs the productive rules of the language, which in turn captures the regularities. This component computes the regularities on-line, utilizing the stored stems from the mental lexicon. (1) is a proposal of what such a rule could look like in relation to the productive rule of the past tense[3]:

(1) V_{past} --> V_{stem} + *d*
(Pinker, 1998, p. 5)

In the case of an irregular lexical entry, there would be stored two forms; the base form and the irregular past tense form. Additionally, there would be grammatical tags which capture the function of the word. (2) captures what such an entry could look like[4]:

(2) V V
Bring --- $Brought_{past}$
(Pinker, 1998, p. 6)

When English is taken into consideration, such an organization seems reasonable. Regular verbs are an open-ended class of verbs seeing as there are thousands of existing words and new ones are created frequently due to the productive nature of the regular paradigm. Irregular verbs, on the other hand, form a relatively small number of verbs (ca. 180), which are formed in idiosyncratic ways (Pinker, 1998, p. 5). Pinker theorizes that the associative memory links features to features, thus capturing semi-regularities among the irregular verbs as described earlier.

> Similar items, which share features, are partly superimposed in the memory representation, allowing the common patterns to reinforce each other, and new items that are similar to learned items will activate the shared features and hence inherit the patterns that have been learned previously, allowing for a kind of generalization. (Pinker, 1998, p. 8)

Since memory has connectionist properties, similar to those of the Rumelhart & McClelland's (1986) model, it is capable of capturing the semi-regular patterns among irregular verbs. Such properties would account for findings indicating that people have the ability to create irregular-sounding novel words (Pinker, 1998, p. 6-9).

Transferring connectionist properties and introducing theoretical modelling of how memory works into the current theoretical framework do cause problems. Introducing these properties accompany the problem of the integration of typed variables in memory (Jackendoff, 2009, p. 164-5). If memory does indeed capture the semiregular patterns among irregular verbs through associative properties, then this begs the question how

[3] Pinker & Ullman (2002) emphasizes, after receiving criticism from Jackendoff (2002), that this is not an actual discrete rule that the theory proses. Rather, it is part of the more general rule of *MERGE* and *UNIFY* applied to constituents. Consequently, the regular formation of the past tense is the unification operation applied to any lexical item with the typed variable V and the past-tense morpheme *-ed.*

[4] Note that this does not take the phonological properties into consideration. Rather, it gives an indication of what such an entry could in principle look like.

typed variables, which enable combinatorial application, are integrated into these memory entries. As Jackendoff (2009) points out in one of his footnotes:

> […] an unadorned network model is not sufficient. As Pinker and colleagues have observed, we need at the very least a notion of a discrete lexical entry - not just a string of sounds – to which syntactic category and meaning can be associated. (Jackendoff, 2009, p. 188)

However, Pinker & Ullman (2002) counter such a view by specifying that the Rumelhart & McClelland-model is not simply glued onto a rule system. Rather, memory bears properties from such a connectionist model, but lexical entries do still have structures relating to semantics, morphology, phonology and syntactic representations. Although this is an interesting theoretical issue, I will not pursue this matter any further due to the limitations of this study.

From the theoretical perspective of a dual-organization of the language faculty, the U-shaped developmental curve is explained in the following manner. To begin with, children have no prior knowledge of the rules of a specific language. Consequently, they memorize each lexical entity as a lexical entry via exposure, including irregular items. This way, they begin producing the irregular forms correctly quite consistently. Through further input, they extrapolate the regular past tense rule, and thus begin marking the regular past tense inflection. Co-occurring with the acquisition of this rule, overregularization errors start taking place. Words that have weaker memory traces due to a lower frequency of exposure are more prone to overregularization.[5] In the absence of a strong memory trace or even in cases where the child has not been exposed to a word, the regular past tense rule acts as a default, resulting in the overapplication of the rule. With time, these errors even out as the memory traces become stronger. During production, the productive rule acts as a default. When producing a past tense verb, one first checks the mental lexicon to see if the entry only contains the stem, or if there is additional information stored such as in example (2) above. In cases of irregularities, the productive rule is blocked when the irregular form is found in the lexical entry, resulting in the successful production of the irregular form. In cases of regular verbs, on the other hand, there are no irregular entries found which can block the rule, and the successful production of a regularly inflected word can thereby take place.

Not only has the dual-mechanism theory been a controversial theory in regard to connectionism. Other factors such as cross-linguistic considerations and the psychological reality of the theory have been contributed to the controversy. For instance, Behrens & Tomasello (1999) question how such a theory which claims universality among humans can account for languages with minimal morphological markings such as Chinese Mandarin. In contrast, there is a large body of work utilizing different sources of data which supports the theory among other languages than English (cf. Clahsen, 1999; Rodriguez-Fornells, Clahsen, Lleó, Zaake, & Münte, 2001; Sonnenstuhl & Huth, 2002). Clahsen's (1999) study of lexical entries and rules in the German inflectional system of both nouns and verbs, suggests a dual-organization of the language faculty. Sonnenstuhl & Huth's (2002) study of processing and representation of German *-n* plurals also supports Clahsen's work. In Clahsen's (1999) multidisciplinary study, he provides

[5] However, this still fails to explain overregularization and production errors of highly frequent words such as *goed* from the verb *to go*. As far as I can tell, Pinker (1998) does not fully explain similar production errors.

evidence for a dual-organization from four different sources, such as child language acquisition and brain event-related potentials (ERP's). The ERP studies referred to by Clahsen (1999) convey that some of his previous work on a morphological violation paradigm have shown a disassociation between regular and irregular verbs and nouns in German through the production of a negative waveform starting at around 200 msec. called a left anterior negativity (henceforth LAN) (Williams, 2018).

> We found the LAN for incorrect irregulars in each of the three participle experiments and even more strikingly in a different inflectional system, noun plurals. In linguistic terms, regularization are violations of affixation […]. A LAN was only found in such cases. Thus, the LAN found under these conditions can be interpreted as reflecting processes involved in morphological structure binding. (Clahsen, 1999, p. 1004)

Interestingly, results from a study of the ERP violation paradigm in Catalan also support Clahsen's findings and conclusions through the replicability of the LAN-effect (Rodriguez-Fornells et al., 2001). The authors of the study conclude:

> From a linguistic perspective, our findings indicate that the division of labour between rule-based and memory-based processes (as posited by dual-mechanism models of morphological processing) applies not only to inflection, but to stem-formation processes such as those found in the Romance languages. (Rodriguez-Fornells et al., 2001, p. 57)

It is important to note, as emphasized by both Rodriguez-Fornells et al. (2001, p. 48) and Clahsen (1999, p. 1002), that the exact functional properties of the LAN are still a controversial topic. It is still striking how this same effect has been replicated in different languages such as Catalan, English, German, and Italian under a morphological violation paradigm, and how cross-linguistic studies seem to support the dual-mechanism theory (Williams, 2018).

There have been two notable studies on U-shaped learning in SLA. The first noteworthy study on U-shaped learning in SLA addressed the acquisition of the Spanish copula contrast *ser* 'to be' and *estar* 'to be' among a native Spanish-speaking group (N=19) and a native English-speaking group of second language learners of Spanish (N=7) (Geeslin and Guijarro-Fuentes, 2006). The age of the participants ranged from 20 to 46 for the native Spanish-speaking group, and 20 to 47 for the native English-speaking group. By providing the two groups with multiple choice contextualized preference tasks, the Spanish-group was only tested once while the English-group was tested on four occasions during a 3 year degree program in Spanish. Using the Spanish-group's responses as the standard, the responses from the English-group were judged accordingly. In sum, Geeslin & Guijarro-Fuentes (2006) found patterns of U-shaped learning in the development of the copula choice for the English-group of second language learners of Spanish. Interestingly, a shaper U-shape curve was evident for a sub-group of the native English-speaking participants who studied abroad in Spain for 4 months.

The second interesting study pertains to lexical acquisition as opposed to grammar acquisition. Shirai (1990) investigated U-shaped learning among Japanese EFL learners of English of three different proficiency groups, along with one group of American native speakers of English. His main claim is that U-shaped learning in L1 acquisition is different from L2 acquisition. Through acceptability judgements, he investigated if the groups exhibited U-shaped behaviour in lexical acquisition. However, no clear picture emerged

from his data. Regarding lexical development, Shirai (1990) concluded that U-shaped behaviour is observable for non-prototypical positive transfer items when the target L2 language is typologically similar to the subject's L1. Furthermore, he wrote "[…] that whether U-shaped behavior is observed in L2 lexical acquisition is highly item-specific and unpredictable, even though I tried to examine some conditions" (Shirai, 1990, p. 6). He also posed the idea of a three-stage model where a transitional period from L1-dependency to L1-independency, intermediated by a restructuring stage, is the cause of U-shaped behaviour. The conditions for this are also task specific. Certain theories concerning linguistic knowledge tied to bilingualism postulates how becoming bilingual leads to a form of structuring and restructuring of language (Grosjean, 1992, p. 57). Gass & Selinker (2008) also support the idea of the restructuring of knowledge and further hypothesize that it is the main cause of the three stages of U-shaped behaviour and learning. Shirai (1990) also seemed to consider the restructuring of knowledge as the main cause of U-shaped behaviour, and how chunk learning is also a contributing factor. One of the limitations of Shirai's study is that it only addresses lexical development, and not grammar development.

2.3 U-shaped learning as a universal vs. L1 transfer

The notion of restructuring knowledge as the main cause of U-shaped learning and development is not radically different from what the dual-mechanism theory proposes. The extrapolation and acquisition of a rule constitute a restructuring of knowledge. It is therefore conceivable that U-shaped learning is not only limited to L1 acquisition, but that it is also affects SLA. The interesting question here is if U-shaped learning is in fact a universal trait of language acquisition, affecting not only L1 acquisition but also L2 acquisition. However, L1 transfer is also another possible explanation.

A case study of Patty (Lardiere, 2003), a Chinese speaker with English as her L2, examined her knowledge of the past and finite tense. The case study, using both spoken and written data, revealed a rather high deletion rate of the past tense morphemes. It turned out that in both written and spoken data combined, Patty marked the regular verbs only 5.8 % of the time, and the irregular verbs 46.08 % of the time (lexical main verbs only). There were similar consistent rates of marking verb inflection when counting all verbs including auxiliaries, modals and copula, respectively 5.8 % regular and 41.3 % irregular. Lardiere (2003) states that this is a case study and that the study is in no way representative of all native Chinese speakers with English as their L2. Lardiere (2003) hypothesizes that Patty's L1 seems to influence her L2 production. Furthermore, she proposes that the English past tense is not acquirable by Patty or other Chinese speakers of English. The study provides evidence of L1 transfer as opposed to the universality of U-shaped learning. One question that needs to be asked, however, is whether or not the data captured Patty's early stages of U-shaped learning. Similar to U-shaped learning exhibited in L1 acquisition, Patty did not seem to have formed the regular past tense rule, which might parallel children in initial stages of acquiring the regular past tense productive rule where they heavily rely on the memorization of items. Compared to the regular verbs, she correctly marked the irregular past tense at a more consistent basis. Furthermore, the error rate is perhaps not very surprising when considering Patty's earliest point of formal exposure to English, which was at age 16 in Hong Kong, China. This is further explained by Long's (1990) finding that there are possibly multiple

sensitive periods in SLA. While reviewing research on maturational constraints on language development, Long (1990) found considerable evidence for a sensitive period for acquiring morphology and syntax in SLA. "Native-like morphology and syntax only seems to be possible for those beginning before age 15" (Long, 1990, p. 280). Johnson & Newport (1989, as cited in Long, 1990, p. 271) support this by unambiguously showing a clear advantage in acquiring morpho-syntactic features before age 15. Patty's performance, in light of this, is therefore perhaps not surprising. Furthermore, whether Patty's performance speaks in favour of L1 transfer or U-shaped learning is an open question.

The Bottleneck Hypothesis, as proposed by Slabakova (2008, as cited in Slabakova, 2016), contends that functional morphology is among the hardest parts to acquire in SLA. Moreover, it is especially difficult to acquire when there is a mismatch present between the L1 and the target L2. Jensen (2016) tested the Bottleneck Hypothesis using a group of Norwegian L2 learners of English between the age ranges of 11-12 and 15-18. It turned out that subject-verb agreement, which has no equivalent in Norwegian, was indeed more challenging for the test subjects. Furthermore, although their proficiency level increased, subject-verb agreement still seemed to be a persistent problem. However, this study focused on the interface between syntax and morphology, whereas this current study focuses more on the relation between morphology and grammar. Norwegian and English are both languages that make a distinction between regular and irregular verbs. Since there are similarities between these two languages in regard to regular and irregular verb inflection, it is therefore conceivable that this aspect of grammar is not as challenging to acquire for Norwegian L2 speakers of English as with other non-similar features, as illustrated by Jensen (2016). However, one interesting question one might raise is whether U-shaped learning, as opposed to simply L1 transfer, is universally true in language acquisition. Phrased differently; do we exhibit a U-shaped developmental curve in acquiring regular and irregular aspects of a grammar system in second language acquisition, or do we simply transfer relevant aspects from our L1 into the target L2?

2.4 Research questions

Given the previous research presented above, I pose the following research questions:

> RQ1: Do Norwegian L2 learners of English exhibit a similar learning curve in relation to the acquisition of irregular and regular aspects of verb morphology as L1 learners of English do?
>
> RQ2: Do we find evidence of dual organization in the L2 lexicon and thus a universal pattern?

To the best of my knowledge, there is a lack of research on the topic of U-shaped learning during L2 acquisition, at least with regards to functional grammar. The purpose of this study is therefore twofold. On the one hand, it is a confirmatory study seeking to explore whether U-shaped learning patterns exist or not within the confines of the three groups that will be tested. On the other hand, it is also an exploratory study which

investigates at what level of accuracy the participants perform, as well as examining how much the participants tend to overregularize irregular verbs.

We hypothesize that there will be an increase in the rate of overregularization errors for group 2 (9th grade) as compared to group 1 (8th grade) and group 3 (10th grade). Although this forms an inverted U-shape, this is still part of the same phenomenon which is referred to as U-shaped learning. The *U* in the term *U-shaped learning* models the dip in accuracy, but the name refers to the phenomenon in its entirety. That is; the different U-shaped curvature, whether they are inverted or not, are associated with the dip in accuracy and are still part of the same phenomenon at hand. In sum, accuracy rates and overregularization rates are linked together, and any results should reflect both aspects simultaneously. We further hypothesize that evidence of the U-shape would be indicative of universality contrary to L1 transfer. Since Norwegian and English have similar regular and irregular inflectional paradigms, where the regular past tense rule is highly productive, it is conceivable that L1 transfer would mainly involve the memorization of the lexical items of English. One would transfer knowledge of the Norwegian productive rule into English which maps rather closely. Consequently, there would be few curves in the rates of overregularization errors, and furthermore no U-shaped learning.

3 Methods

In order to investigate whether or not a U-shaped learning curve does occur among Norwegian L2 learners of English in the acquisition of irregular and regular aspects of verb morphology, a cross-sectional design was chosen for this experimental quantitative study. Due to the confines of a master's thesis, we chose to mainly focus on irregular verbs. U-shaped learning involves a regression of accuracy in the production of irregular words over a certain duration of time. Time is therefore one of the important independent variables when examining U-shaped behaviour. In order to accommodate for this in the cross-sectional design, grade/proficiency level was therefore chosen as a proxy for age.

Level of accuracy in conjunction with overregularization errors of irregular verbs were additionally of significant interest when observing U-shaped behaviour. During L1 acquisition, co-occurring with the regression of accuracy, is an inverted U-shape in the rate of overregularization errors which peaks as accuracy dips. Consequently, the rate of overregularization errors across time is therefore especially interesting.

Data from elicitation tasks modelled after the classic Wug-test design (Berko, 1958) were collected from participants at three different school-grade levels, respectively in the 8th, 9th, and 10th grade. Data from each group were collected via an online form and analysed statistically afterwards. In the following section, a description of the participants in this study is provided, in addition to the materials used and the procedure itself. Finally, after a section addressing the analysis, the validity of this study is discussed.

3.1 Participants

The participants in this study were students at a lower-secondary school in Trondheim, Norway. They were recruited randomly from within three different grades, respectively in the 8th grade (N=20), 9th grade (N=24), and 10th grade (N=23). However, participants with known learning deficits or other relevant developmental deficits were excluded. In addition, participants with a different L1 than Norwegian, and those performing beneath a certain accuracy level (readdressed later in 3.4) were also excluded. The participant count in the final analysis are as follows in the 8th grade (N=17), 9th grade (N=19), and 10th grade (N=15).

Since there were participants in the 8th and 9th grade who were younger than 16, consent had to be given by their parents in addition to by the participants themselves. The parents were given a detailed information and consent sheet which offered brief information about the project, data protection and handling, and their legal rights during the project. A separate information and consent sheet, with less legal information, were given to the participants themselves, which in turn gave a brief outline of what they were

expected to do whilst filling in the online form. Participants in the 10th grade who were 16 years old were given a specially dedicated information and consent sheet since they could legally provide consent to participate themselves. A questionnaire related to the participant's language background, as well as their medical and developmental background, was also given to the parents and filled out by them. The participants were informed of this in their information sheet. This did not apply to the participants in the 10th grade who instead received all the relevant information themselves. See appendix for more information on the information sheets and consent sheets.

3.1.1 The students

Students in Norway start learning English formally during their first year of attending school. After completing the 7th grade and upon entering the lower-secondary level, they will have received 366 hours of education on English, whereas during their time in a lower-secondary school, they will receive an additional 228 hours (UDIR, 2013). Normally, students in the 8th grade are between 13 and 14 years, 14 and 15 years in the 9th grade, and 15 and 16 years in the 10th grade.

After the research project was approved by both the school and the Norwegian Centre for Research Data (henceforth NSD), a select group of volunteer teachers helped grant access to their classes. Sensitive information relating to the students' health were also collected in order to uncover factors that could potentially influence their language development.

3.1.2 The parents

Before any testing took place, consent had to be given first by the parents before the participants themselves could give any consent to participate. Since socio-economic factors seems to be an important confound, data were also collected on the parents' socio-economic background. In doing so, a questionnaire was handed out along with the information and consent sheet. This mainly focused on their educational background and occupations. For an overview, see Appendix B. Due to the scope of this thesis, I will not go into the details of this data.

3.2 Materials

3.2.1 Stimuli and elicitation tasks

The participants' knowledge of grammatical aspects of irregular verbs were of main interest for this study. Subjects were therefore mainly tested on verb production. Additionally, U-shaped behaviour has been observed for irregular nouns as summarized in chapter 2. Participants were tested in 43 irregular verbs and 10 irregular nouns.

In order to create the elicitation tasks, I began by obtaining an overview of the learning material used by the school for teaching English. There were three books which the

school used for each grade. According to one of the teachers involved with this project, these books were frequently used as part of the learning material. First, a list of the irregular verbs found across all books in the books' glossary was established. It amounted to 93 irregular verbs. A short list of 14 irregular nouns was also created in a similar fashion, with the exception of a few words. For instance, instead of using the word *salmon*, I chose *fish* seeing how it occurs more frequently.

As presented in the literature review, frequency of exposure seems to be an important factor in relation to U-shaped learning and the production of overregularization errors. Subsequently, I created two conditions for verbs and two conditions for nouns based on frequency ratings; an easy condition and a hard condition. In order to do this, I checked the frequency ratings for both the verbs and the nouns in BYU's *iWeb: The intelligent Web-based Corpus*. An open search for verbs, grouped by lemmas, resulted in a list of the top 1000 most frequently occurring verbs in the corpus. The same process applied to the nouns. Of the 93 irregular verbs found in the books, 20 with a frequency ranking of smaller than 50 out of the 1000 hits (RAW FREQ greater than 6056787) were chosen for the easy condition, whilst an additional 20 with a frequency ranking greater than 300 of the 1000 hits (RAW FREQ smaller than 1013581) were chosen for the hard condition of verbs. Of the 14 irregular nouns found in the books, 5 with a frequency ranking of smaller than 100 out of the 1000 hits (RAW FREQ greater than 4027114) were chosen for the easy condition of nouns, whereas an additional 5 with a frequency ranking greater than 600 of the 1000 hits (RAW FREQ smaller than 1193738) were chosen for the hard condition of nouns. An additional 3 verb items with a RAW FREQ smaller than 10566 were also added to the list of items to be tested. We assumed that the participants were highly unlikely to successfully produce the correct past tense forms of these items. These three items were added to see what types of mistakes the participants would make. In sum, they were given 43 tasks related to irregular inflection of verbs, and 10 tasks related to irregular inflection of nouns.

After the four (five if you count the unlikely condition) lists were created to form frequency-based conditions, elicitation tasks based on Berko's (1958) Wug-test were created afterwards. The following example (3) illustrates what the tasks, meant to test verb knowledge, looked like:

(3) STEAL: Paul likes to steal things.
However, during the party, Paul _____ Lisa's heart.

The target word is specified in its non-finite form for the verbs (spelled out in capital letters to highlight that this is the target word) and is further used in the present tense in an example sentence. In the second sentence, a blank is left where the participant is asked to fill in the irregular past tense form of the word according to context, which in the above example would have been *stole*. As for the elicitation tasks for the nouns, they were designed in a similar fashion. (4) shows an example of how they were formulated:

(4) MAN: A baseball team cannot consist of only one man.
A baseball team consist of at least 9 _____.

3.2.2 Questionaire

Two separate online forms were created to accommodate for the participants who needed parental consent and those who were able to give consent themselves. The online form intended for the participants in the 8th and 9th grade included a question related to known diagnoses pertaining to language impairment or learning impairment. As a requirement from NSD, this question was excluded from the online form intended for the 16-year old participants for them to be legally allowed to give consent themselves. Both forms were identical aside from this one question.

3.2.3 Electronic survey platform

SurveyGizmo was chosen as the online survey platform to collect data on the responses to the elicitation tasks, as well as personal habits surrounding the participant's use of English. To maintain the anonymity of each participant and their responses, all candidates were issued a personal code which corresponded to their name on a physical key sheet. Participants were also asked to fill in their date of birth as an extra precaution in case he or she had written down the wrong personalized code. If a participant were to withdraw consent, cross-checking the date of birth with the consent sheet, along with the personal code, would make for a more reliable way of identifying the correct participant. For the purpose of analysing the data, reliably matching the participant with information on relevant diagnoses provided by the parents was important in cases where I would have to exclude the subject form further analysis.

SurveyGizmo also collected reaction times on each question. These measurements gave information on the amount of time each participant spent on answering each task. It might be interesting to see whether there are any patterns in the amount of time spent on answering each task, in addition to checking whether they spend more time in cases where they produce errors.

12 background questions inspired by Strætkvern's (2017) and Grønning's (2016) background information were added into the electronic survey. These questions focused on the participant's language background (both first and second language), self-rated level of proficiency, language usage, and exposure to English.

3.3 Procedure

After the project was designed, it was submitted to NSD for evaluation before being accepted. Prior to any testing, consent sheets were given to both parents (see appendix B) before the participants received any information (see appendix A). Ahead of each testing session, a list of participants was created from the consent sheets that had been collected by the involved teachers. Each testing session took about 60 minutes and was held at the school during teaching hours. Some of them took place in a mixed classroom with other non-participating students. Other sessions were held in a room separate from the other non-participating students. During mixed classroom sessions, non-participants

were given individual work by the teacher and encouraged to keep quiet so to not disturb the participating students. All testing sessions were administered by this thesis' author.

3.4 Analysis

Although accuracy rates over time are important in relation to U-shaped learning during L1 acquisition, they are likely not quite the same in L2 acquisition. Importantly, children during L1 acquisition are exposed to very similar types of input. In L2 acquisition on the other hand, the input varies a lot more and has arguably a greater impact on the language development than in the L1. This will necessarily impact the individual's vocabulary size, as well as how they store words. Thus, we are not expecting accuracy rates alone to be a good measure of U-shaped learning in L2 acquisition. It is conceivable that accuracy will in fact increase with age and level of proficiency. Rather, we expect accuracy rates in conjunction with overregularization over time to be a better way of measuring this phenomenon in the L2.

Overregularization errors might be a better predictor to U-shaped learning during L2 acquisition. As previously mentioned, overregularization errors increase for a duration of time during the dip in performance, as observed during the development of the L1. Furthermore, this forms an inverted U-shape for the production of such errors. We are therefore interested in measuring the progression of overregularization errors in the L2 to see to what degree it follows L1 development. Ultimately, we expect to see curvature in the rates of overregularization errors across the groups.

Given the abovementioned rational, grade as a predictor of overregularization for mainly verbs, but also for nouns to a certain extent, is of special interest. For the analysis, grade was used as the independent variable, whereas accuracy and overregularization were used as dependent variables. All responses to the elicitation tasks were first compiled in Microsoft Excel where they were followingly judged according to three categories; accurate responses, overregularization errors, and other types of errors[6]. Percentages were calculated for each variable. For each variable, the mean score and standard deviation were calculated for each group. The distributions for each group were also check in the form of histograms and boxplots.

Looking at the distributions in the rate of accuracy within each group lead to further exclusion of participants on the basis of developmental deficits, underperformance (under 50% correct answers overall verb production), as well as bilingualism.

Participants were tested for 43 verbs and 10 nouns. For the verbs, as mentioned above, three conditions were created; an easy condition, a hard condition, and an unlikely condition. Note that the unlikely condition, consisting of only three items, was excluded from the analysis in order to create a balanced analysis between highly frequent items (items in the easy condition) and less frequent items (items in the hard condition). Only the easy and the hard condition, which consisted of a balanced 20 highly frequent items

[6] Responses that were a combination of approximation (changed stem vowel) + regular past tense suffix were added to the category of 'other types of errors'. These are typically judged as normal overregularization errors in L1 production. However, these were marginal cases for all three groups, and were added to *other errors* since they are a combination of two types of errors.

and 20 less frequent items, were used for the analysis. As previously mentioned, the *unlikely condition* was simply used to see what types of mistakes the participants would make. However, due to the word limit of this thesis, I could not pursue an analysis of these items.

The results focus on the relation between grade and rate of accuracy, as well as grade and rate of overregularization errors. A final selection of relevant boxplots and other figures are presented. The main focus is on overregularization errors. Closer examination of the distributions within each variable revealed how they were not normally distributed. ANOVA and t-tests could therefore not be performed, and a Wilcoxon test was consequently adopted to test certain relations that were of interest.

3.5 Validity

A major benefit of using elicitation tasks is that it enables us to test specific aspects of the participant's knowledge of words, such as tense, regular vs. irregular inflectional morphology, and noun count for instance. The tasks restrict the type of answers a participant can provide, which has certain benefits. Seeing as the target word is specified and the context of the sentence predicts a single accurate answer to a greater extent, evaluating the data normatively and accurately is therefore straightforward. The method additionally restricts the amount of information for further statistical processing. Finally, it is a well-established method applied within linguistic study disciplines.

The validity of this study is based on certain assumptions. One central assumption is that the participants undergo a similar learning curve, which in turn is the rational for the cross-sectional design. Furthermore, we are assuming that the students in the selected grade interval (8th, 9th, and 10th grade) are still developing their metalinguistic awareness. A master's thesis by Evensen (2014) demonstrates how metalinguistic awareness is under development for Norwegian children and adolescents learning English as an L2 between the ages 10 to 15 years. On a critical note, it is possible that the age interval within these three grades is too compact for L2 acquisition. This could mean that a larger interval is needed to see any real curvature for either accuracy or overregularization rates. It is also possible that we would be able to see more of a dip in the production rate of overregularization errors by testing students at the end of upper secondary (VG3).

Since reaction times were recorded in a classroom setting with varied levels of control, they cannot be considered 100% reliable. Participants would sometimes, even though they were encouraged not to do so, get up during the experiment and leave the computer for a few seconds. This affects the reliability of the data seeing how some of the measurements are inaccurate. Sample size is another factor since the groups are neither equally big nor more than 20 participants. It would have been preferable to have a greater number of participants, especially in the third group.

As a final point, there is an added benefit of reliability when utilizing an online platform to distribute elicitation tasks. An online platform used to administer the tasks minimizes the effect a researcher might have on the participating subject and the responses to each

task. Furthermore, it ensures that all the participants receive the tasks by means of the exact same procedure.

4 Results

The following section presents the results from the data collection. Accuracy rates for verbs and nouns are presented descriptively. The subsequent sub-section focuses on the rate of overregularization for overall verb and noun production, before providing finer details on the differences between the easy and the hard condition for verbs and nouns. Overregularization rates were further processed in R for statistical inferencing. Descriptive statistical analyses illustrate information on the reaction times for different types of responses. Finally, information on the other types of errors made when producing verbs and nouns are also reported. For reference, participants in group 1, 2 and 3 correspond respectively to 8th grade, 9th grade, and 10th grade.

4.1 Accuracy rates

A closer inspection of the data revealed that there were a few participants that performed beneath a 50% accuracy rate. Initially, participants with known dyslexia had not been excluded since there seems to be an open question as to what extent dyslexia affects the L2. Although it is not a language impairment, dyslexia is a developmental deficit nonetheless. We know that children with this diagnosis struggle in Norwegian schools. However, we do not know the exact reason behind this. Regardless, three participants with this known diagnosis were excluded from further analyses. An additional 4 participants were also excluded from further analyses due to poor performance (beneath 50% accuracy), and balanced bilingualism. The remaining number of participants were respectively in 8th grade (N=17), 9th grade (N=19), and 10th grade (N=15).

Figure 1 to 4 show the data after the abovementioned outliers were excluded:

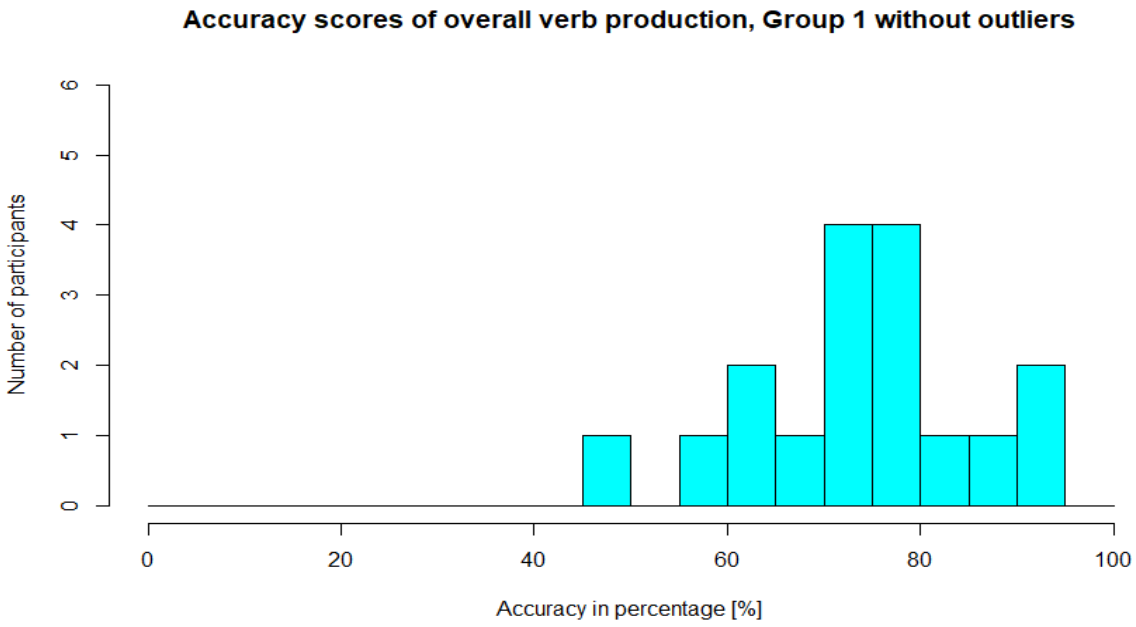


Figure 1: Overall accuracy of verbs in group 1.

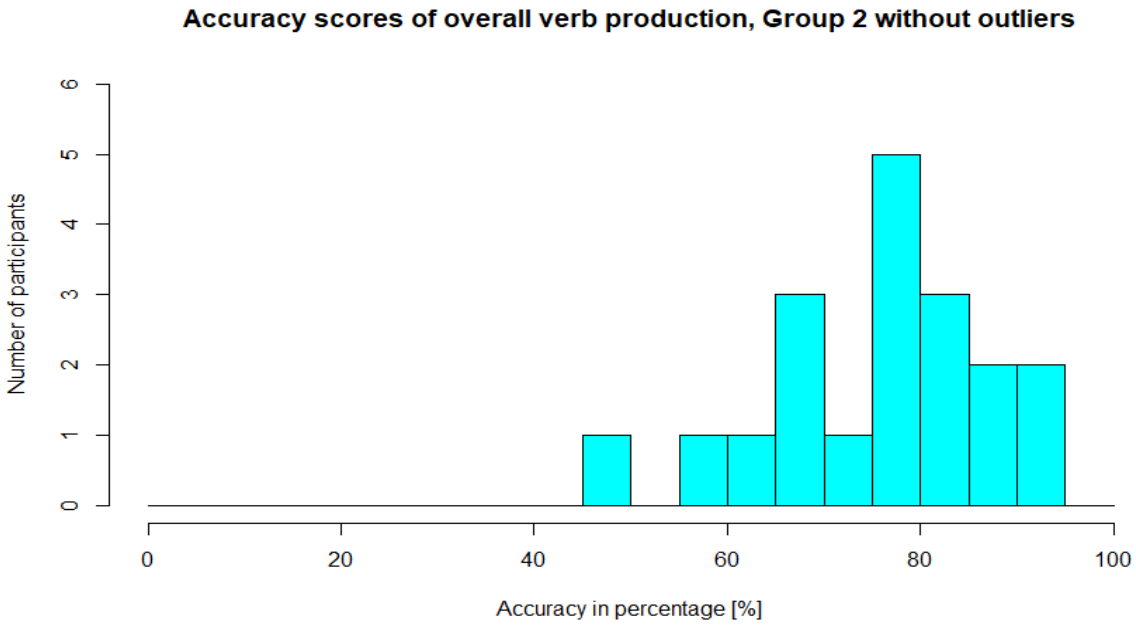


Figure 2: Overall accuracy of verbs in group 2.

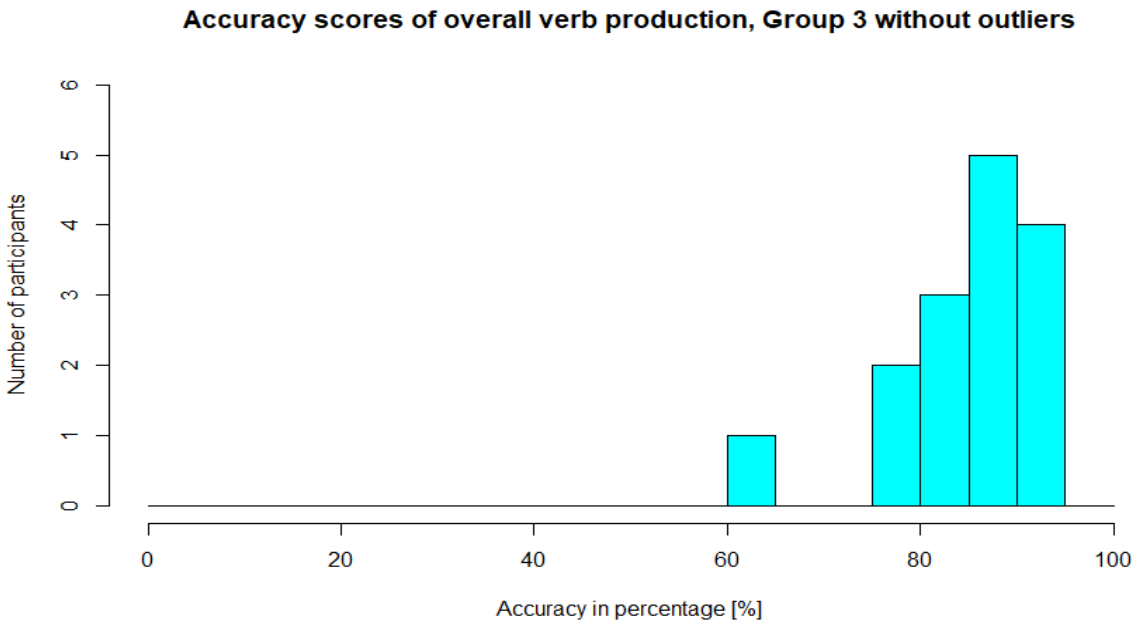


Figure 3: Overall accuracy of verbs in group 3.

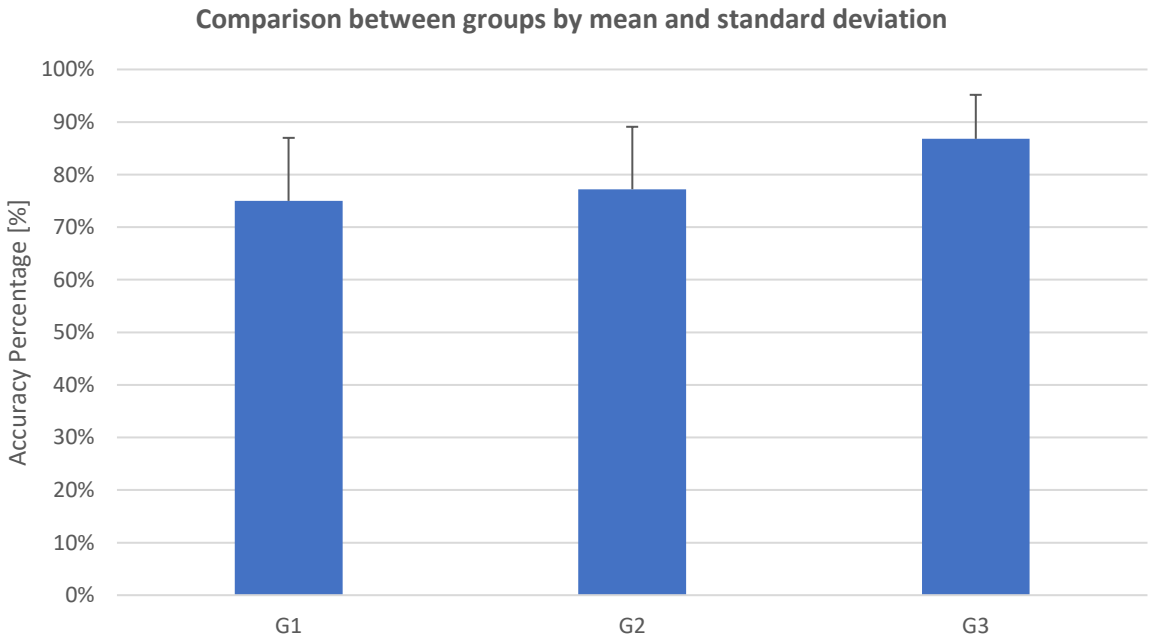


Figure 4: Comparison between groups by mean and standard deviation for overall accuracy of verbs.

Whereas figure 1 to 3 show the distributions among the groups, figure 4 compares the mean average performance of each group and includes standard deviation error bars. As expected, overall accuracy rates seem to slightly increase with grade and age, at least in regard to the overall verb production. Note that the variability also decreases with grade.

As for overall accuracy for nouns, figure 5 provides data on the distribution within each group, as well as comparisons between them:

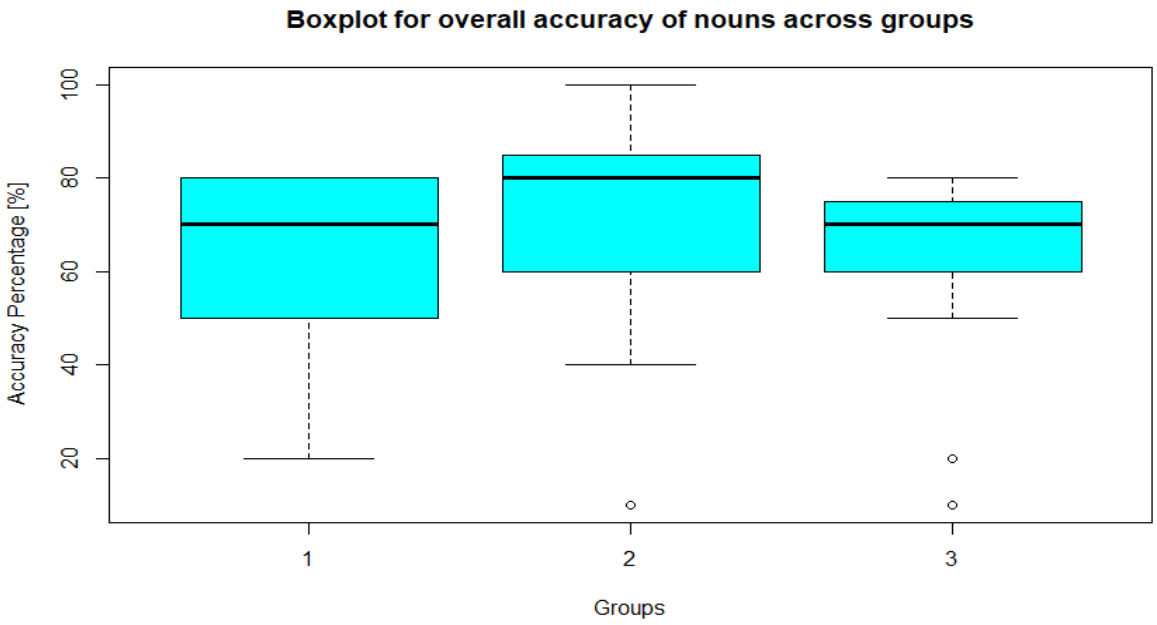


Figure 5: Comparison between groups for overall accuracy of nouns.

Although performance seems to increase with grade for verb production, this does not seem to be the case for the overall noun production (figure 5). The mean varies between 61.3% and 70.5%. Accuracy rates between the groups for nouns seem consequently to be more stable. Note that the variation in accuracy decreases between group 2 and 3. Additionally, there are a few outliers still present for group 2 and 3.

By comparing the easy condition with the hard condition for verbs within each group, figure 6 demonstrates how the participants are sensitive to frequency effects. Consider the following figure:

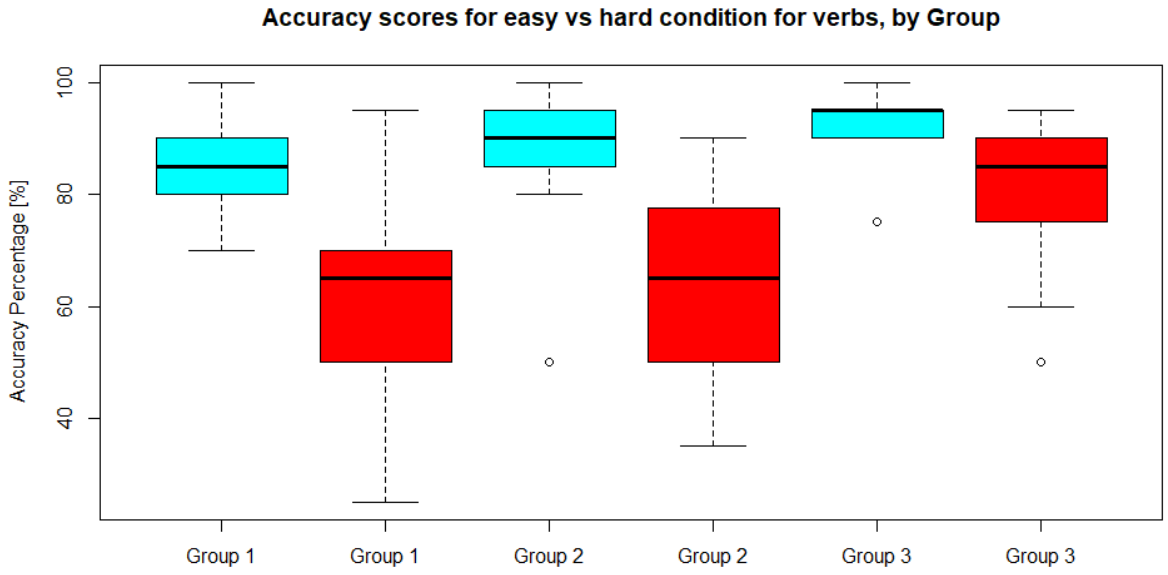


Figure 6: Accuracy scores for easy (turquoise) vs hard (red) condition for verbs, by Group.

There are clear differences between each condition within each group, which indicate that the participants are sensitive to frequency effects. Performance appears to increase in the hard condition across the three groups. Additionally, there is greater variation in the accuracy scores in the hard condition as opposed to the easy condition across all three groups. While accuracy increases with grade, variation within both conditions decreases.

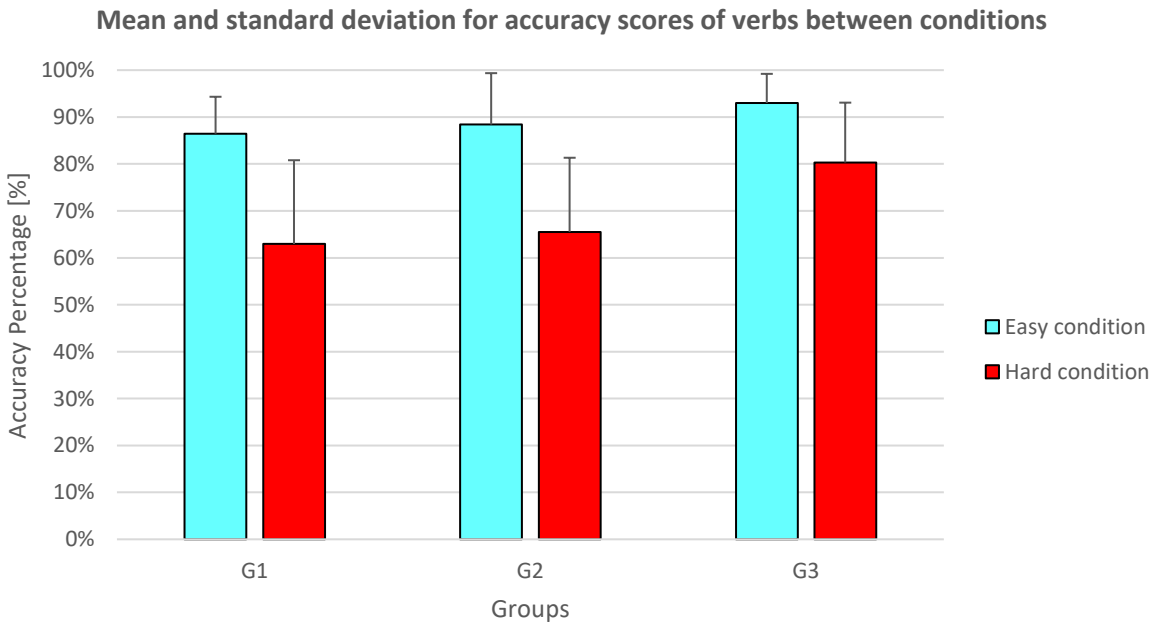


Figure 7: Comparison between groups for mean scores and standard deviation, easy and hard condition of verbs.

Figure 7 tells us the following information. On average, participants in group 1 were accurate 86.4% of the time for the highly frequent verbs in the easy condition, but only 62.9% accurate for the less frequent verbs in the hard condition. For the second group, they exhibited an accuracy rate of 88.4% for highly frequent verbs, and 65.5% for the less frequent verbs. Finally, in group 3, participants were 93% accurate for frequent verbs, and 80.3% accurate for the less frequent ones.

Sensitivity to item frequency are also visible for nouns, as illustrated below:

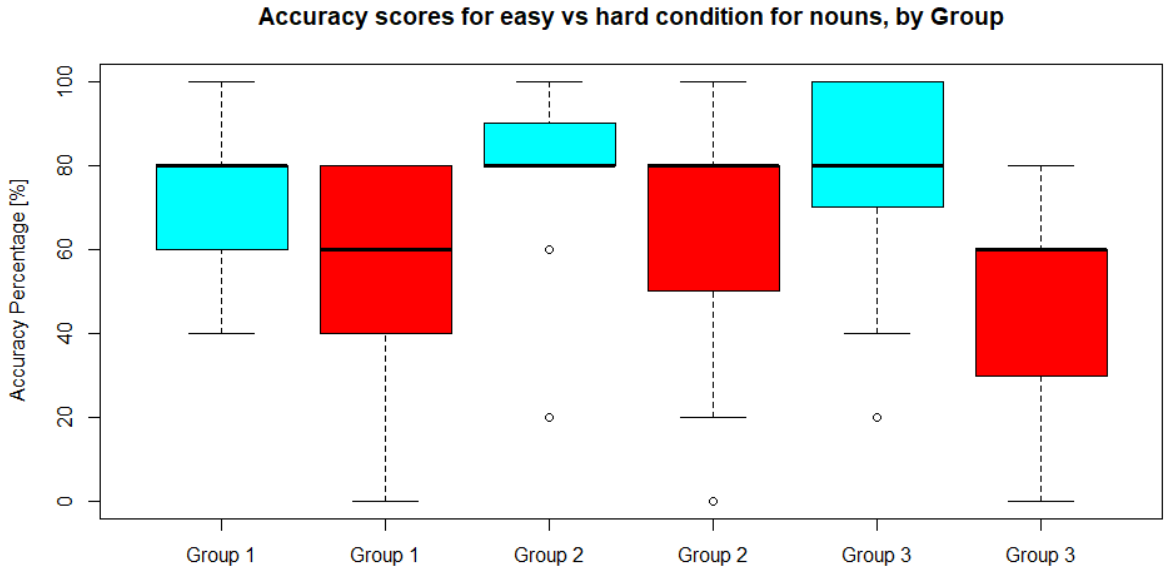


Figure 8: Accuracy scores for easy vs hard condition for nouns, by Group.

The distributions within each variable are more skewed for the nouns than for the verbs (compare figure 8 to figure 6). Contrasting conditions within groups, we see a similar trend where there is less variation in accuracy for the easy condition than the hard condition. We still see some outliers in group 2 and 3.

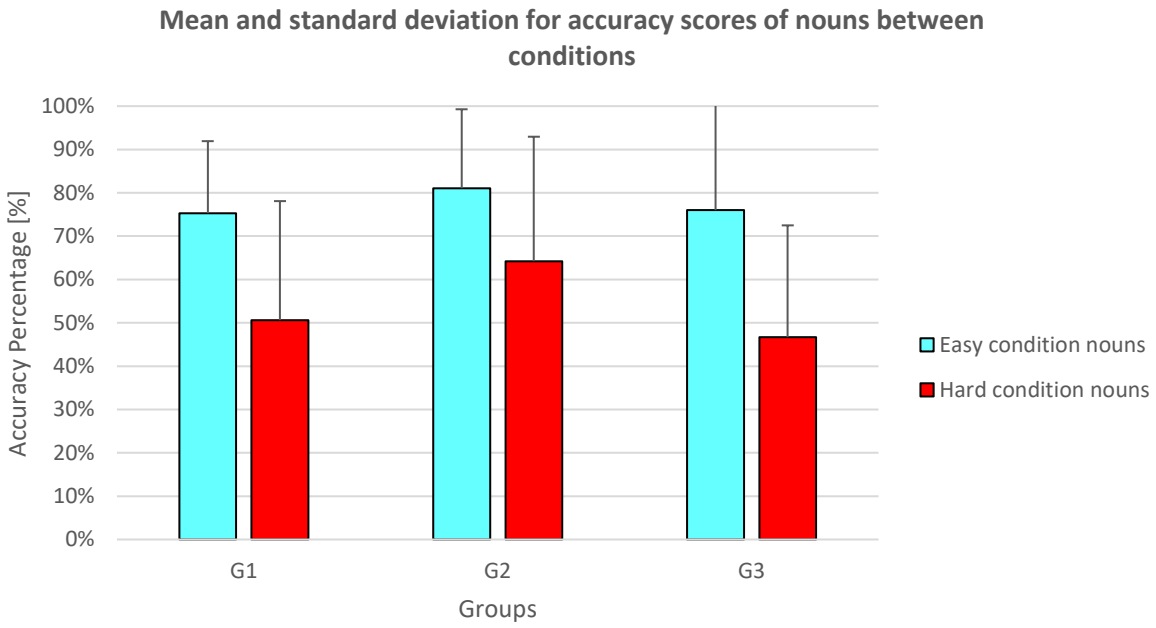


Figure 9: Comparison between groups for mean scores and standard deviation, easy and hard condition of nouns.

In the first group, participants produced highly frequent irregular nouns correctly on average 75.2% of the time, but only 50.5% of the time for the less frequent irregular nouns. As for the second group, they were accurate 81% and 64.2% of the time for highly frequent verbs and less frequent verbs respectively. Finally, in the third group, accuracy rates were 76% and 46.6% by the same comparisons. Considering the contrast in the average performance within each group, frequency effects are observable.

Turning to the inferential statistics, the following tests were performed in order to examine whether the differences in the accuracy were significantly different. As shown in several of the previous boxplots, many of the distributions appeared skewed. A Shapiro-Wilk test was performed to give further feedback on this observation (Table 1):

Shapiro-Wilk test for normality, verb accuracy			
	Group:	**W:**	**P-value:**
Overall rate of accuracy	G1	0.90045	p-value = 0.09669
	G2	0.9411	p-value = 0.276
	G3	0.83475	p-value = 0.01064
Easy condition for rate of accuracy	G1	0.96375	p-value = 0.7027
	G2	0.73515	p-value = 0.0001505
	G3	0.79585	p-value = 0.003254
Hard condition for rate of accuracy	G1	0.94704	p-value = 0.4114
	G2	0.94991	p-value = 0.3938
	G3	0.90045	p-value = 0.09669

Table 1: Shapiro-Wilk test for normality, accuracy distributions for verbs.

As indicated by table 1, quite a few of the dependent variables were not normally distributed since the p-values were less than 0.05. If a p-value is less than 0.05, then the null-hypothesis, which is that the data is normally distributed, is rejected and the alternative hypothesis, that the data is not normally distributed, is accepted. For overall accuracy rates, group 2 distributed normally, whereas group 1 and 3 did not. For the easy condition, group 2 and 3 did not distribute normally, while group 1's responses did. Finally, all the responses within all the groups distributed normally for the hard condition.

Since some the data did not distribute normally, the non-parametric unpaired version of the Wilcoxon test[7], which does not assume a normally distributed population in the samples, was applied in R to see whether the distributions overlapped or not. Table 2 provides the details of the applied test and group comparisons:

[7] Also referred to as the Mann-Whitney U test.

Wilcoxon test between groups for overall accuracy		
Groups being compared	**W:**	**P-value:**
Group 1 and Group 2	W = 137	p-value = 0.4454
Group 2 and Group 3	W = 66	p-value = 0.008083
Group 1 and Group 3	W = 48	p-value = 0.002734
Wilcoxon test between groups for accuracy, easy condition		
Groups being compared	**W:**	**P-value:**
Group 1 and Group 2	W = 123.5	p-value = 0.2246
Group 2 and Group 3	W = 98	p-value = 0.1133
Group 1 and Group 3	W = 62	p-value = 0.01185
Wilcoxon test between groups for accuracy, hard condition		
Groups being compared	**W:**	**P-value:**
Group 1 and Group 2	W = 144.5	p-value = 0.5979
Group 2 and Group 3	W = 65	p-value = 0.007152
Group 1 and Group 3	W = 53	p-value = 0.004987

Table 2: Results from group comparisons on accuracy rates. Result of an unpaired Wilcoxon test.

The NULL-hypothesis, that there are no significant differences between sample A and B, is proven if the p-value is greater than 0.05. The test gives an indication of whether the two distributions (A and B) overlap or not, and if they differ significantly. For overall accuracy, as well as for accuracy in the easy and hard condition, the distributions for group 1 and 2 did not differ significantly. The distributions, when comparing group 2 and 3 for accuracy in the easy condition, did not differ either. However, we did see a significant difference between group 2 and 3, as well as between group 1 and 3 for overall accuracy in verb production. The same is true for the same comparisons in the accuracy rates in the hard condition. As for the accuracy in the easy condition, the only significant difference was between group 1 and 3. In sum, we see that group 1 and 2 never differ. Group 2 and 3 differ sometimes, whereas group 1 and 3 differ systematically.

4.2 Overregularization rates

Figure 10 gives an overview of the overall overregularization rate for the irregular verbs. It also illustrates the distribution within each group, whereas figure 11 compares the mean averages of overregularization rates between them.

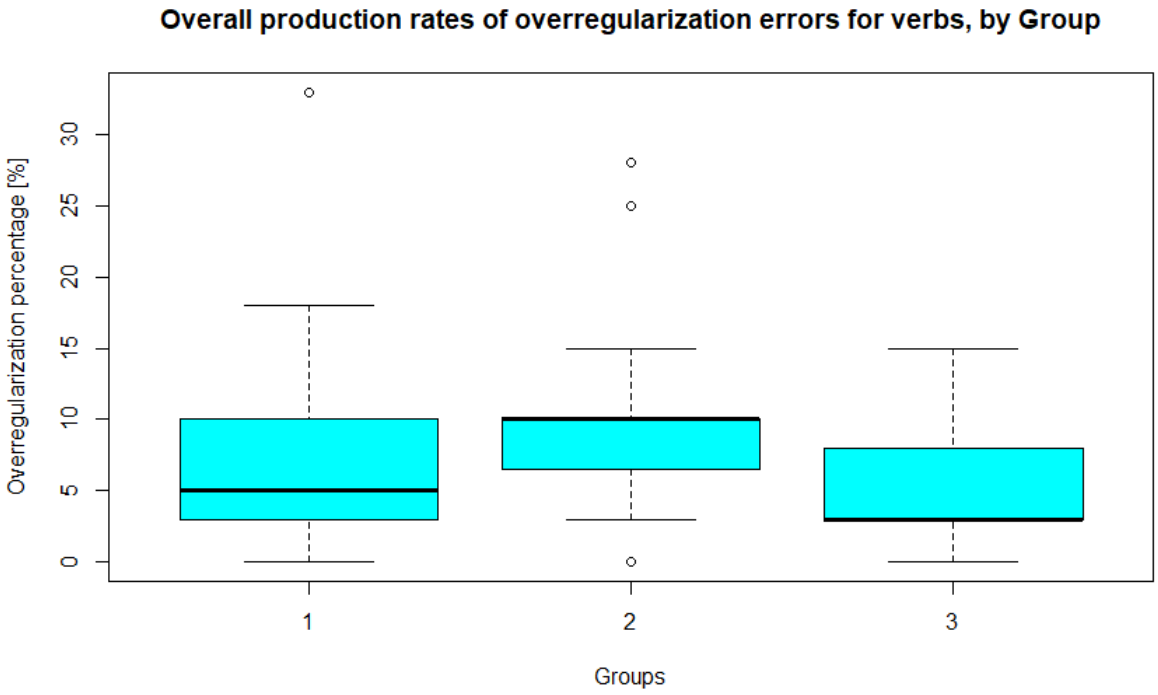


Figure 10: Overall production rates of overregularization errors for verbs, by Group.

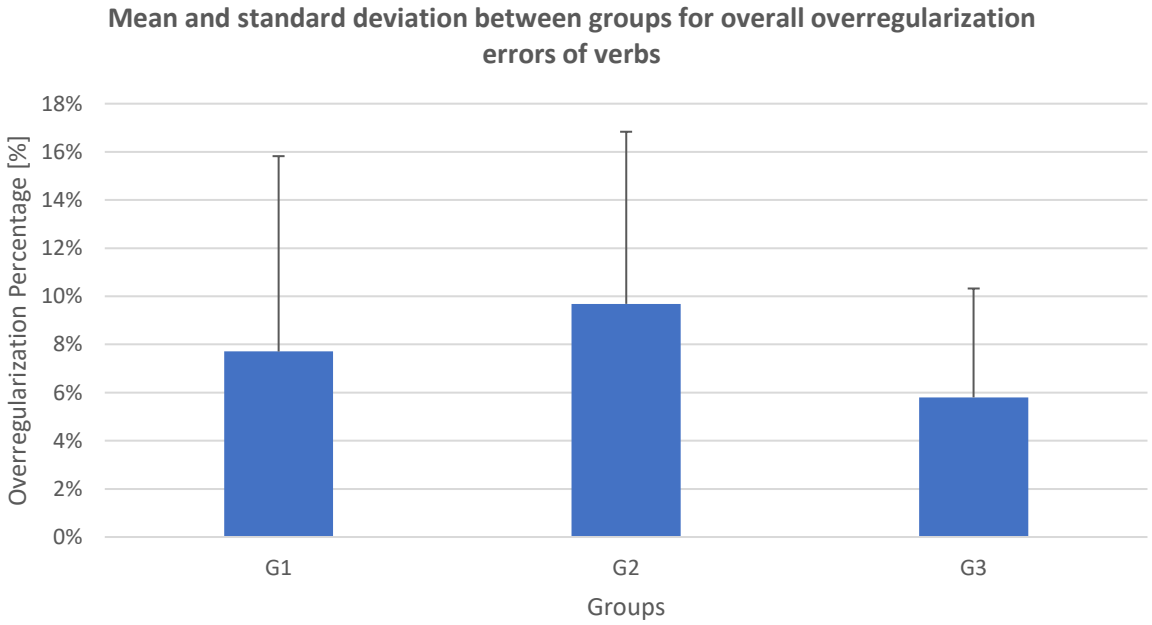


Figure 11: Comparison between groups for overall production rate of overregularization errors for verbs.

Overall overregularization rates seem to be similar across groups, ranging from 5.8-9.6% on average. The distributions, however, look skewed (Figure 10). Furthermore, there are four outliers present. Conversely, note that the medians across the groups form an inverted U-shape which peaks in the 9th grade.

The figures below provide an overview of the overall overregularization rate of irregular nouns. Consider figure 12 and 13:

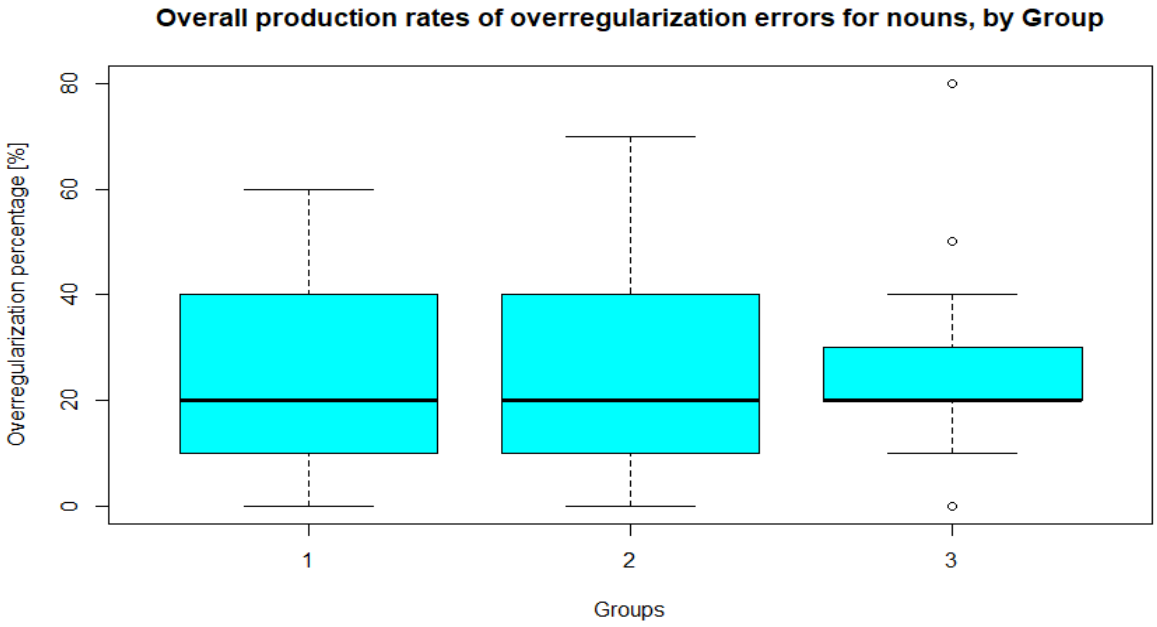


Figure 12: Overall production rates of overregularization errors for nouns, by Group.

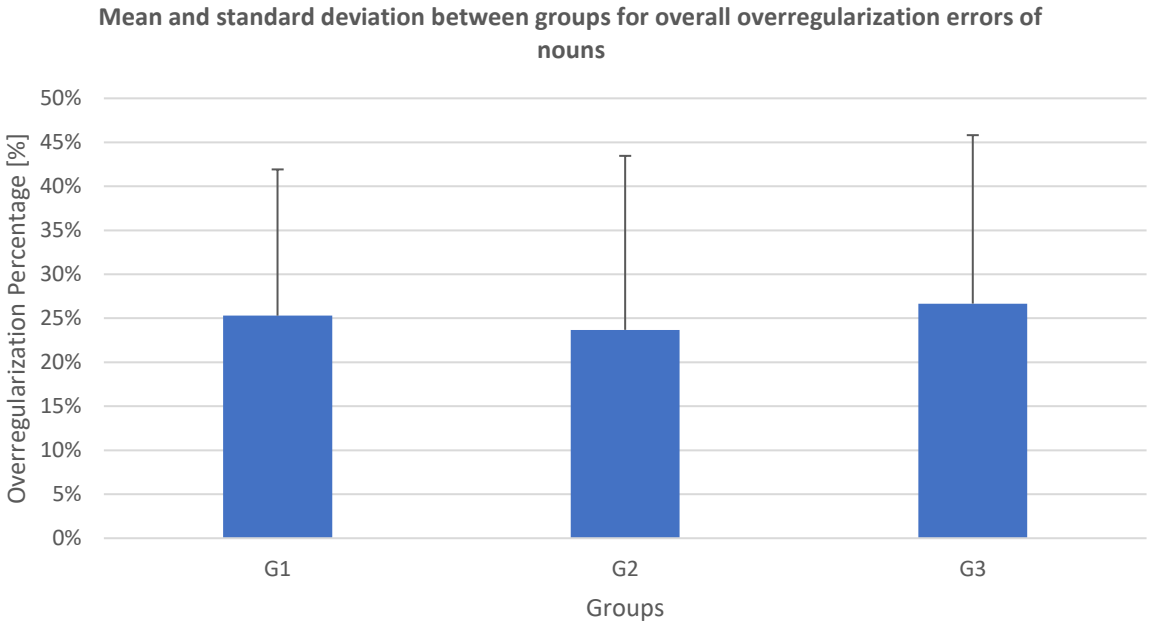


Figure 13: Comparison between groups for overregularization errors of nouns.

The groups performed surprisingly similar. There is no change in the level of overregularization errors for nouns across groups. It is noteworthy that the variation groups more tightly together for group 3 as opposed to the two other ones (Figure 12).

When contrasting the easy condition with the hard condition for verbs within each group, some interesting differences begin to emerge. Consider the following comparisons within each group between conditions (Figure 14 and 15):

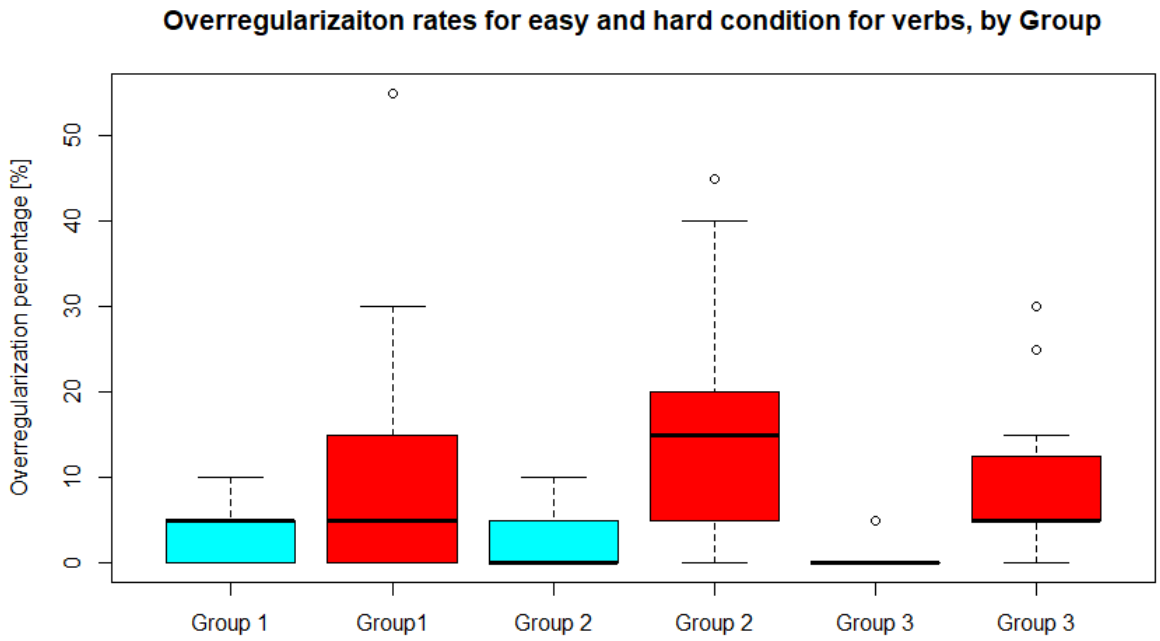


Figure 14: Comparison between easy condition (turquoise) and hard condition (red) for rate of overregularization of verbs, by Group.

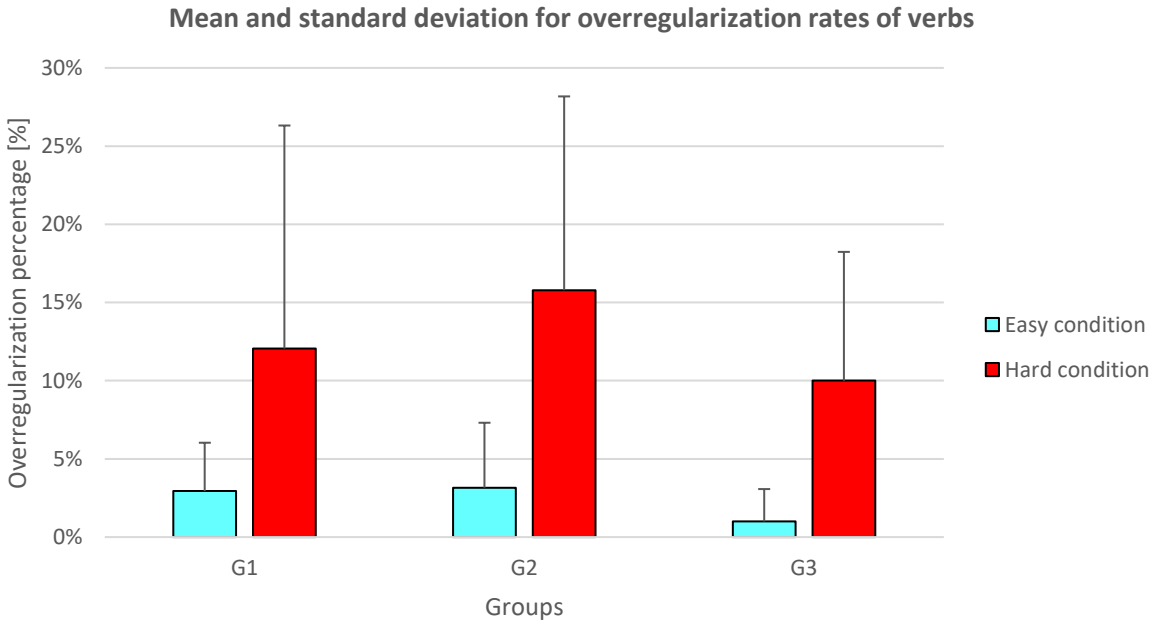


Figure 15: Comparison between groups for mean scores and standard deviation, overregularization of verbs, easy and hard condition.

Although the rate of overregularization errors seems to be stable between group 1 and 2 within the easy condition of verbs, it seems to drop significantly between group 2 and 3 (see Figure 14). Group 3 exhibits a floor effect for this kind of error for highly frequent verbs. Furthermore, figure 14 indicates that rate of overregularization within the easy condition for group 3 is tightly distributed, and that there is a lot less variation among the participants. Altogether, every group produced more overregularization errors for less frequent verbs when compared to the highly frequent ones. We also see that the rate of overregularization errors in the hard condition rises in group 2, thus forming an inverted U-shape.

A similar trend is not as prevalent when observing the overregularization rates of the irregular nouns. Consider figure 16 and 17:

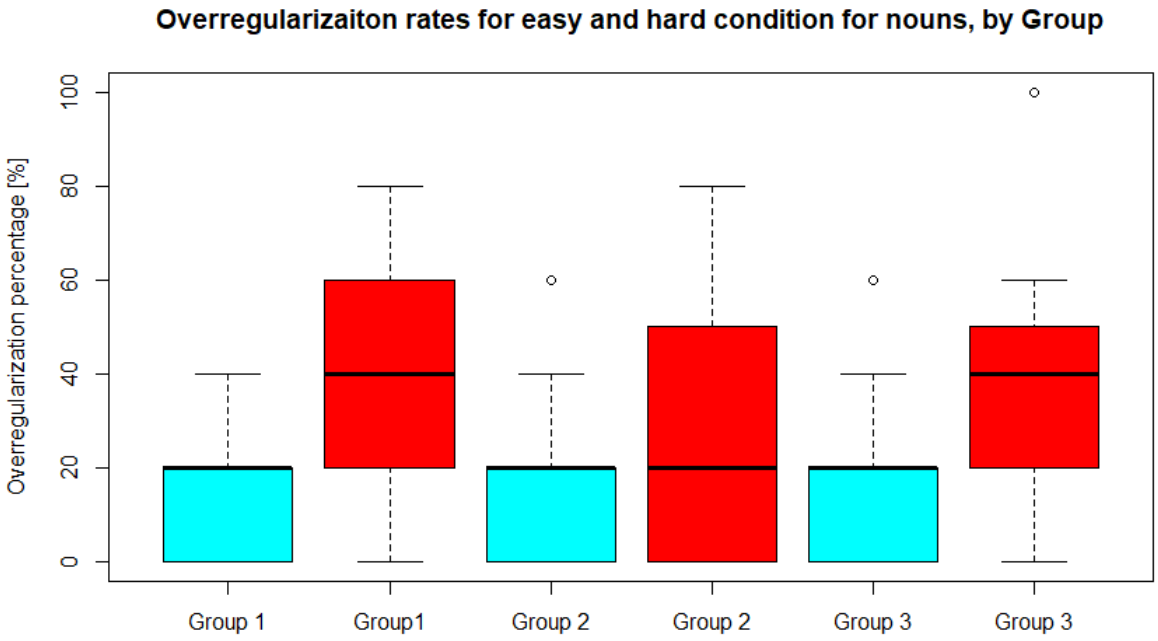


Figure 16: Comparison between easy condition (turquoise) and hard condition (red) for rate of overregularization of nouns, by Group.

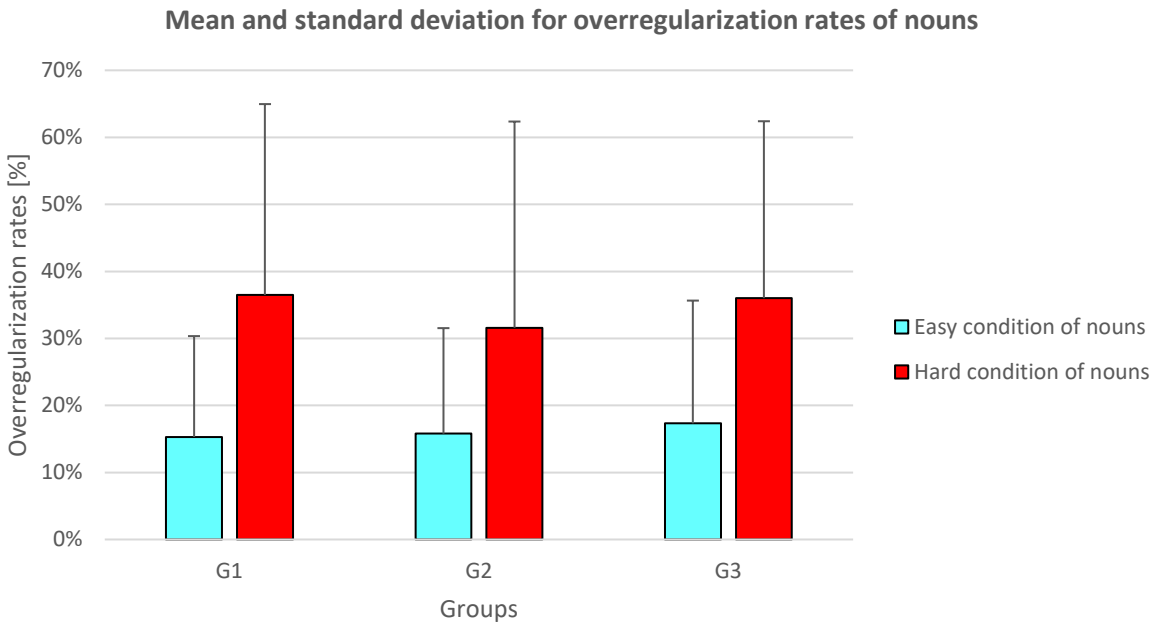


Figure 17: Comparison between groups for overregularization rate of nouns, easy and hard condition.

Overregularization rates of nouns appear to be stable across the three groups in the easy condition. There is an overall higher error rate in the hard condition contrary to the easy condition. Interestingly, we see a U-shape when looking at the medians across the groups for the hard condition (figure 16).

Turning to the inferential statistics, the following tests were performed in order to examine whether the differences in the rates of overregularization errors of verbs were statistically significant. By observing the data in figure 14, it led us to believe that the error rates were not normally distributed. A Shapiro-Wilk test for normality was performed and gave the following results (table 3):

Shapiro-Wilk test for normality			
	Group:	**W:**	**P-value:**
Overall rate of overregularization	G1	0.78067	0.001116
	G2	0.85953	0.009626
	G3	0.81933	0.00658
Easy condition for rate of overregularization	G1	0.75245	0.000485
	G2	0.71361	8.11E-05
	G3	0.49944	3.48E-06
Hard condition for rate of overregularization	G1	0.78956	0.001464
	G2	0.90579	0.06199
	G3	0.81184	0.005237

Table 3: Shapiro-Wilk test for normality, overregularization distributions.

One out of three groups did distribute normally within the hard condition whereas none were normally distributed in the other conditions where the P-values were smaller than 0.05.

Consequently, the Wilcoxon test was applied in R to see whether the distributions overlapped or not. The test was applied between the groups to see whether there were any significant differences between the distributions within each condition. The test gave the following result (table 4):

Wilcoxon test between groups for overall overregularization rate		
Groups being compared:	**W:**	**P-value:**
Group 1 and Group 2	W = 127	p-value = 0.2719
Group 2 and Group 3	W = 195	p-value = 0.06649
Group 1 and Group 3	W = 141	p-value = 0.6142
Wilcoxon test between groups for overregularization rate in the easy condition		
Groups being compared:	**W:**	**P-value:**
Group 1 and Group 2	W = 165	p-value = 0.9161
Group 2 and Group 3	W = 180	p-value = 0.1198
Group 1 and Group 3	W = 171	p-value = 0.05497
Wilcoxon test between groups for overregularization rate in the hard condition		
Groups being compared:	**W:**	**P-value:**
Group 1 and Group 2	W = 120	p-value = 0.1875
Group 2 and Group 3	W = 181.5	p-value = 0.1721
Group 1 and Group 3	W = 124	p-value = 0.9073

Table 4: Wilcoxon test between groups for overregularization rates.

The distributions overlapped and did not differ significantly when comparing group 1 and 2 (p-value = 0.2719) for the overall overregularization rates, as well as for group 1 and 3 (p-value = 0.6142). Interestingly, the distributions are close to being significantly different when comparing group 2 and 3 (p-value = 0.06649) for the overall production rates. As for the overregularization rates in the easy condition, the p-value for the distributions of group 1 and 3 (p-value = 0.05497) is also close to being significant. However, it is still not below the threshold of 0.05. On the other hand, one could consider it to be a borderline significant p-value. Note that none of the tests were able to give exact p-values since there were ties in the data, which means that all the p-values are approximate computations. In sum, none of the distributions, whether tested together or separated by condition, differed significantly.

An additional Pearson correlation test was performed to see whether the participants' self-rated level of proficiency had any correlation with the rate of overregularization errors. No correlation was found between the participants' self-rated level of proficiency and the overall rate of overregularization errors of verbs (as shown in table 5)

Self-rated level of proficiency against:	Group:	Correlation:
Overall rate of overregularization	Group 1	0.00171426
	Group 2	-0.2487454
	Group 3	-0.4861471
Easy condition for rate of overregularization	Group 1	-0.2362087
	Group 2	-0.3860624
	Group 3	-0.559017
Hard condition for rate of overregularization	Group 1	0.05121768
	Group 2	-0.1553049
	Group 3	-0.4214636

Table 5: Pearson correlation test between overall rate of overregularization for verbs and the participant's self-rated level of proficiency.

4.3 Reaction times on accurate and overregularized answers

As described in the previous chapter, reaction times were also recorded for each response. Figure 18 and 19 comparers the reaction times on accurate and overregularized responses respectively:

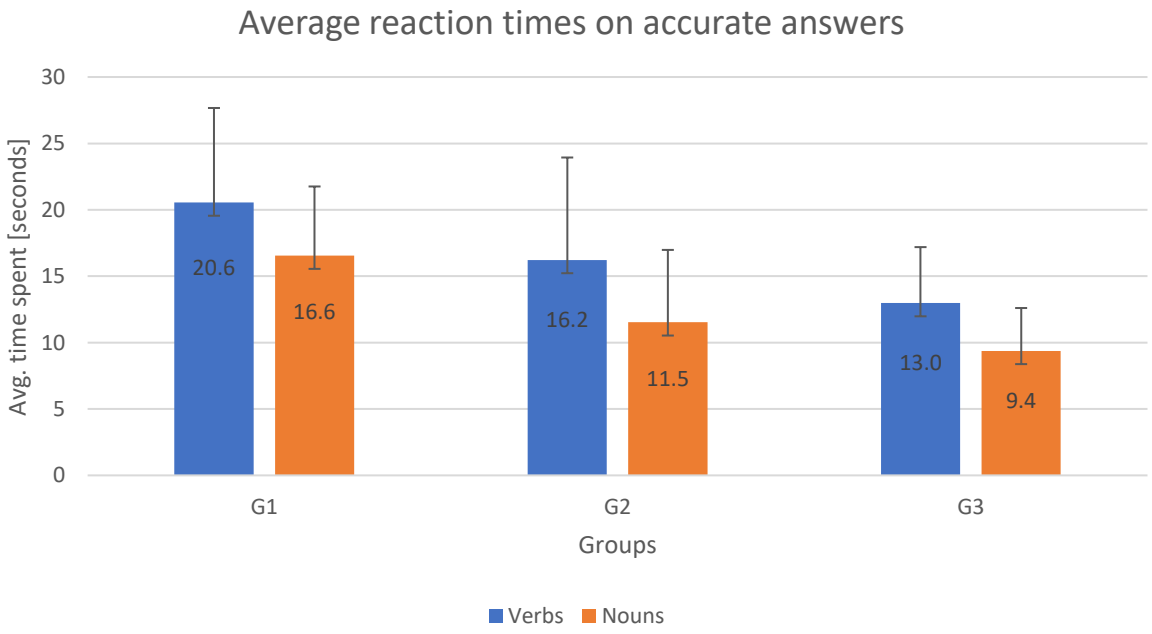


Figure 18: Average time spent on accurate answers with standard deviation error bars. Average time spent on accurate verb responses (blue) and for nouns (orange) are shown.

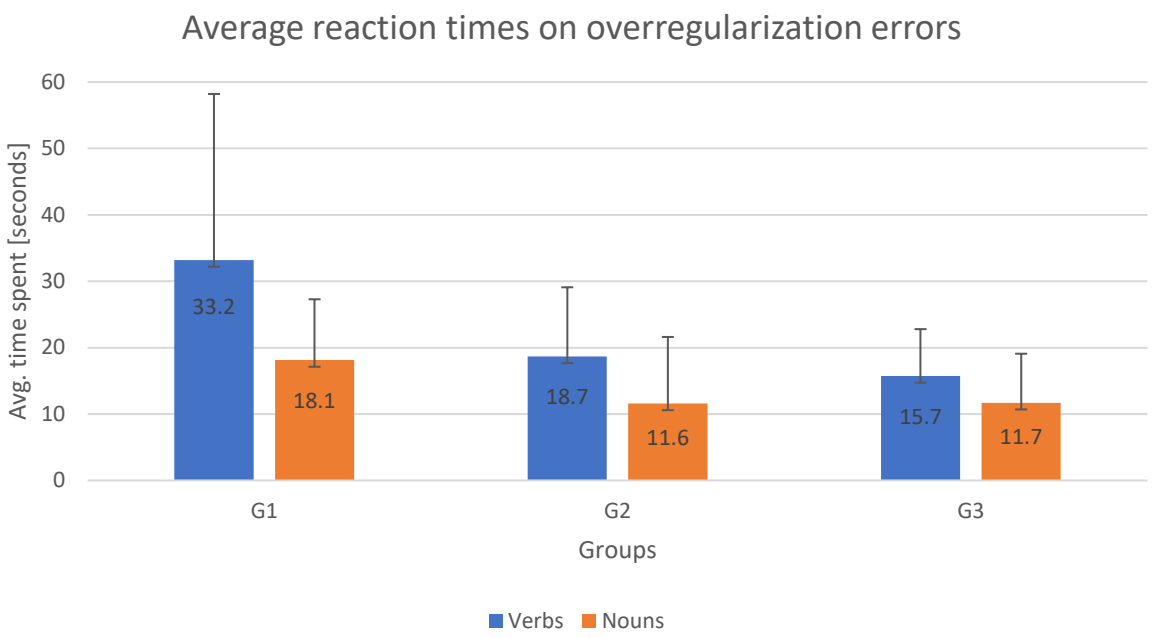


Figure 19: Average time spent on overregularization errors, including error bars for standard deviation. Average time on overregularization errors for verbs (blue) and for nouns (orange).

As figure 18 illustrates, reaction times on accurate answers for both verbs and nouns seem to decrease with grade and level of proficiency. A similar trend is observable for overregularization errors for both verbs and nouns (Figure 19). As indicated by the standard deviation error bars, the variability does also seem to decrease with grade for time spent on verbs and nouns in both figures. Average reaction times for overregularization errors for nouns seem to flatten out between group 2 and 3.

4.4 Other errors types in relation to verbs

All the responses were judged according to three categories, and the errors were sorted into two categories when evaluating the responses to the tasks. Error category 2 consisted of overgeneralization errors. Error category 1 encompassed all other forms of errors. Although often considered overregularization errors, cases of approximation plus the regular inflectional suffix *-ed*, such as *frozed* for the past tense of *to freeze*, were added together with all the other errors (separate from Error category 2). However, they were counted separately seeing as there was only a small amount of this type of error. Presented in this sub-section are examples of the other forms of errors produced by the participants in this study.

Error category 1 included the following; bare forms, avoidances, omissions, phonological approximations, and wrong tense. Additionally, it included cases of approximation plus overregularization. Following are examples of each kind of error:

(3) "SHINE: Brian likes it when the sun shines on him.
Yesterday, the sun _shine____ through the clouds on Brian."

(4) "SEE: Paul loves to see movies on his spare time.
Last night, Paul __watched____ Mad Max: Fury Road at home."

(5) "TELL: Paul likes to tell jokes at parties.
At yesterday's party, Paul __?____ a funny joke that made everybody laugh!"

(6) "CLING: When rock climbing, Brian clings tightly onto the rocks.
Yesterday, Brian __clong____ tightly to the rocks so he would not fall."

(7) "SET: Every year, this guy sets out on a new journey.
What did he do last year?
He __sets___out on a journey."

(8) "FREEZE: Brian likes to prepare all his food a week in advance, and freeze it afterwards.
Last night, after preparing his food, Brian _frozed__ his food."

Example (3) illustrates a typical case where the participant provides the bare form as opposed to the fully inflected past tense form. (4) is a case of avoidance where the participant provided a different lexical word than the target word (as indicated in capital letters). (5) is a case of omission, whereas (6) shows how the participant has made a phonological approximation of the target word *clung*. Example (7) is a case where the present tense, as opposed to the past tense, was applied. Finally, (8) illustrates a case of an approximation plus overregularization of an irregular verb.

Seeing as bare forms and cases of approximation plus overregularization are quite interesting, table 6 offers further information about the bare forms and table 7 does so for cases of approximation plus overregularization.

	Sum:	**Mean:**	**Median:**
8th grade	23	1.35	1
9th grade	32	1.68	1
10th grade	11	0.73	1

Table 6: Bare forms for verbs.

	Sum:	**Mean:**	**Median:**
8th grade	3	0.17	0
9th grade	1	0.05	0
10th grade	0	0	0

Table 7: Cases of approximation + overregularization for verbs.

Bare forms, as in bare stems, are often attributed to children's early L1 language production. It has also been noted as a common error made by bilingual school-aged children (Jacobson & Yu, 2018). Although participants overall seem to continue to produce bare forms across the groups (median = 1 as seen in Table 6), the production rate of such errors seems to drop substantially from 9th grade to 10th grade.

Interestingly, we find an inverted U-shape in the amount of bare forms across the three grades as seen in table 6.

4.5 Other error made in relation to nouns

Same categorization process as applied to the verbs was applied to the nouns. Beneath are examples of the types of errors found in the data which were; bare form (9), avoidance (10), omission (11), phonological approximation (12), and approximation plus overregularization (13).

(9) "LIFE: A regular person has one life.
A cat, on the other hand, has 9 __life__".

(10) "GOOSE: Today, I only saw 1 goose.
Yesterday, I saw a flock of __ghosts____."

(11) "OX: Farmer Bob has 1 ox.
Farmer Brendon has 5 ___?___.!"

(12) "GOOSE: Today, I only saw 1 goose.
Yesterday, I saw a flock of __gise____."

(13) "GOOSE: Today, I only saw 1 goose.
Yesterday, I saw a flock of __geeses____."

Table 8 compares the amount of bare forms between the three groups, whereas table 9 does so for cases of approximation plus overregularization.

	Sum:	Mean:	Median:
8th grade	8	0.47	0
9th grade	5	0.25	0
10th grade	5	0.33	0

Table 8: Bare forms for nouns.

	Sum:	Mean:	Median:
8th grade	2	0.1	0
9th grade	3	0.15	0
10th grade	1	0.06	0

Table 9: Cases of approximation + overregularization for nouns.

In the case of nouns, bare forms were singular nouns in a plural noun context. Interestingly, there were fewer bare forms for nouns than for verbs (compare Table 8 with Table 6). This could be due to fewer elicitation tasks dedicated to nouns. Regardless, bare forms for nouns do not seem to be a common error for the participants in this study.

5 Discussion

The research questions, as proposed in section 2.4., were as follows. 1) Do Norwegian L2 learners of English exhibit a similar learning curve in relation to the acquisition of irregular and regular aspects of verb morphology as L1 learners of English do? 2) Do we find evidence of dual organization in the L2 lexicon and thus a universal pattern?

Prior to testing, we expected to find significant differences in the levels of accuracy and overregularization errors when comparing the three groups in this study, in addition to possibly finding ceiling effects. We proposed that evidence for U-shaped learning would speak in favour of the universality of the phenomenon as opposed to simply L1 transfer. The overall picture that emerges after analysing the data seems to indicate that we are observing a part of the U-shaped curve which parallels that which has been observed in L1 acquisition. Furthermore, the results seem to model the later stages of the phenomenon.

5.1 Evidence for U-shaped learning in the L2

The first question in this study sought to determine whether Norwegian L2 learners of English exhibit U-shaped learning in a similar fashion to how L1 learners of English do when acquiring irregular aspects of verb morphology in English. This study provides suggestive evidence indicating that we are in fact seeing a small interval from the later stages of U-shaped learning during L2 acquisition among Norwegian learners of English. During the later stages of U-shaped learning, the theory predicts how the level of accuracy should steadily increase while at the same time co-occurring with a decrease in the rate of overregularization errors. The observations in chapter 4 model this to a certain extent. There were some significant differences between the groups for accuracy rates of verb production, and we found a floor effect for the rate of overregularization errors of highly frequent verb items for group 3. By looking at the statistics for accuracy rates and overregularization rates of irregular verbs and nouns, several observations were made and are summarized as follows.

Firstly, overall accuracy rates improve with each grade, rising from 75% accurate verb responses to 86.8% on average. This improvement is most clearly visible for verb production across the groups. The Wilcoxon test revealed how the distributions, when comparing group 1 and group 3 for the overall accuracy rates, as well when sorted by condition, systematically differed to a significant extent (see table 4). Additionally, the variation in accuracy decreases with grade. As for noun production, a clear rise in accuracy was not visible seeing as it seems to have stagnated (see Figure 9). Conversely, the variability in the rate of accuracy for overall noun production decreases with grade as shown in the same figure. This decrease in variability in conjunction with an increase in verb production performance might be indicative of an increase in proficiency among participants across grades.

Secondly, participants were sensitive to frequency effects. There was a clear frequency effect between highly frequent items (easy condition) and the low frequency ones (hard condition). This effect was visible for accuracy rates for verbs (Figure 10) and nouns (Figure 12). Similarly, the same was true for the rates of overregularization errors for verbs (Figure 18) and nouns (Figure 20). This observation supports Pinker's (1998) idea that the frequency of exposure has an impact on memory traces and consequently the lexical retrieval for words. A higher frequency of exposure leads to stronger memory traces, which means that they are less prone overregularization errors and more prone to being correctly retrieved from the mental lexicon. As described in the methods, the four conditions based on frequency were designed by checking the frequency of each item in BYU's *iWeb*-corpus, which is a corpus consisting of ca. 14 billion words from 22 billion webpages. Seeing as children and adolescents spend increasingly more and more time connected with social media and the internet, checking the frequency of the target words against a corpus based on webpages seemed advantageous. In line with this assumption, the frequency effects support this idea.

Thirdly, there were some observable curves between the groups for overregularization rates in the overall verb production (Figure 14). We additionally saw a U-shape across the groups when looking at the medians in the hard condition for nouns in the overregularization rates (Figure 16). Some interesting differences became evident when differentiating between conditions. In the easy condition for verbs, a floor effect was observed for group 3 (Figure 18). As for the hard condition of verbs, we saw a rise in overregularization rates when comparing the median between the groups which in turn resembled an inverted U-shape. Contrary to expectation, the Wilcoxon test applied to all three conditions (overall, easy condition of verbs, hard condition of verbs as seen in table 4) between the groups indicated no significant differences between the three distributions in each tested condition. An explanation of this result will be discussed in further detail in sub-section 5.3 for limitations. Nonetheless, the floor effect was rather interesting demonstrating how the 10th graders almost did not overregularize highly frequent verb items.

Fourthly, regarding reaction times for accurate answers and overregularized items decrease with grade. Not only does the average time spent on these types of responses decrease with grade, but the variability also decreases with grade. This is consistent for both verb and noun production. Additionally, participants spent more time within each grade on overregularized responses than for accurate responses. This is yet further evidence suggesting an overall increase in proficiency from grade to grade.

Fifthly, a Pearson correlation-test between participants' self-rated level of proficiency and the rate of overregularization errors did not correlate (Table 5). If anything, the correlation between these two factors seemed to decrease with grade. Therefore, the participants' self-rated level of proficiency is not a good predictor of the rate of overregularization within each group.

Finally, the amount of bare forms decreases for verbs and nouns from 8th to 10th grade. Regarding the bare forms of verbs, they increased from 8th grade to 9th grade before decreasing in the 10th grade, thereby forming an inverted U-shape. Bare forms, at least in L1 productions, have been ascribed to early stages of language acquisition. There were more bare forms for the verb production than for the nouns. This could be due to the

amount of elicitation tasks dedicated to testing verb knowledge. They were tested on 43 verb items as opposed to the 10 noun items. This finding could suggest that there is more of a development taking place in regard to verbs than for nouns, which is consistent with the accuracy rates which show more of a progression towards higher performance for verbs.

Comparing overregularization rates, a less clear picture emerges. Rates of overregularization errors in the easy condition of verbs in conjunction with the accuracy rates suggest that we are seeing a small interval of the later stages of the U-shaped curve. Also observed was a floor effect in group 3 for overregularization errors of highly frequent verb items, as well as a borderline significant difference between group 1 and 3. On the other hand, the hard condition of verbs exhibited an inverted U-shape curvature when looking at the medians across the groups. This could suggest that we are observing the mid-section of U-shaped learning. However, the Wilcoxon test showed no significant differences between the three distributions within this condition (hard condition of verbs). It is therefore possible that we are observing the later stages of U-shaped learning as opposed to the mid-section. In line with the hypothesis, accuracy rates in conjunction with overregularization rates seem to be a good measure of the curve. Consequently, the results better model the later stages of the phenomenon.

We have reason to believe that our data speak in favour of U-shaped learning as opposed to positive L1 transfer. Mentioned previously, Shirai (1990) found positive transfer for items that were typologically similar between the L1 and the target L2, at least in regard to lexical acquisition. Transfer, either as negative or positive transfer, is often attributed to L2 acquisition. Our data support the contrary. The first indication of this is connected to the frequency effects that we found throughout several of our measurements when differentiating them by frequency-based conditions. The participants' accuracy is measured in response to the frequency of the forms in the target language English. Moreover, the results show how they are tuned in to the English item frequencies as opposed to transferring the frequencies of the Norwegian items. As grade and proficiency increase, the variability in all measurements decrease and the participants begin to produce language more uniformly. This parallels the development among L1 speakers from adolescent to adulthood. The second indication that positive L1 transfer is not the case is linked to the item-by-item correspondence between Norwegian and English. Although the two languages are typologically similar in their division between a productive past tense rule and sub-regular patterns among irregular verbs, the items used in this study do not have a one-to-one match between the two languages. Phrased differently, some of the items which are irregular in English are in fact regular in Norwegian. Positive transfer would mean that they transfer their existing knowledge of a given verb from Norwegian into English. In the easy condition, 15 out of the 20 items match between the two languages in being irregular, whereas 10 out of the 20 match for the hard condition. Given this, it is conceivable that there would be more persistent errors. In our data, on the other hand, we see more curvature in the rate of overregularization errors as opposed to a stagnant rate of overregularization errors. It would have been interesting to do a closer analysis looking at which items the participants do in fact tend to overregularize to see if the mismatching items are more susceptible to overregularization than the matching ones. However, due to the limitations of this study, I will refrain from delving further into this matter. In sum, if transfer was indeed the case, then it is quite possible that there would be less curvature and no U-shaped learning behaviour since they simply transfer the relevant similarities into the

target language from the L1. However, further studies are needed to confirm this claim. For instance, looking at two languages that are typologically dissimilar in regard to this aspect of English, such as Chinese Mandarin and English for instance, would be advantageous.

5.2 Supportive evidence for a dual-mechanism in the L2?

In answering the second research question, I would first like emphasize that there is no clear answer to whether there is a dual-mechanism organization of the mental faculty in the L2. Firstly, it is unclear whether the type of data that we have can demonstrate a dual-organization of the L2. The elicitation tasks, designed to reflect participant's knowledge of specific word items, only tested irregular verbs and nouns. Conversely, since the participants were not tested on regular verb items, we do not have any information on this sample's knowledge surrounding regularly inflected verb and noun items. Therefore, we do not have a clear picture of the knowledge on the relationship between these two, and we can only say something about the participants' knowledge of irregular word items. Secondly, it still seems contestable whether a dual-organization or a connectionist conceptualization of the mental faculty is indeed the psychological reality.

Both theories have their limitations and shortcomings. For instance, connectionism relies a great deal on the performance of their models, meaning that the theory's plausibility is limited by them to a great extent. Certain connectionist models, such as Plunkett & Marchman's (1991) model, have problems when reproducing a U-shaped learning curve which is similar to observable data from L1 acquisition. As criticized by Jackendoff (2002), the connectionist models do not incorporate variables, which are problematic both to incorporate and to try to live without. Ellis (2003) theorizes how the connectionist models should in theory function similarly in the L2. As noted by Pinker (1991, p. 531), the connectionist models, at least the earlier ones as cited in the research chapter, do not actually store words as separate lexical entries distinct from phonological and semantic content. This in turn raises questions as to how one can differentiate between the two lexicons, and ultimately the two languages.

The dual-mechanism has also been criticized. Due to the word limitation, I cannot go into further detail. For further reference, see McClelland & Patterson (2002), and Behrens & Tomasello (1999). Nonetheless, as presented in chapter 2, there is a large amount of evidence which supports the dual-mechanism theory. If dual-mechanism is indeed a plausible theory of the psychological reality of the organization of the language faculty, then it should also be able to account for not only data from L1 acquisition, but also data from L2 acquisition. One question raised by this study is therefore how L2 acquisition fits into the dual-mechanism theory.

We have reason to think we have indicative evidence for a dual-mechanism organization seeing how we might have found a U-shaped learning curve in the L2. Even though there have been connectionist models which have been capable of modelling the U-shape, it is questionable whether U-shaped learning is supportive evidence for connectionism. Connectionist models utilize the frequency of items to adjust the weight of the nodes in the network, thus in a sense adjusting the strength of the memory patterns. Consequently, it is conceivable that a frequency effect should be the only effect that we

would find in our data, where the most frequent items are reproduced correctly. Furthermore, the theory and the models do not make a distinction between storage and a form of regular mechanism. One simply acquires the correct form through exposure, thereby adjusting the weight of the nodes in the network the more frequently you encounter them. It is therefore rather strange that we find overregularization errors in our data for irregular verbs that occur frequently in English. If we were in fact learning the regular pattern (such as the past tense *-ed* for verbs) that are rule-like, in the form of lexical storage which is a form of memory, as proposed by connectionism, then we should not be seeing overregularization errors for highly frequent items.

The explanation laid forth by the dual-mechanism theory better explains why these errors occur in the first place. In learning the patterns in a language, we rely on memory during the early stages. As development continues, we discover the rules of the language which begin to interfere with the memorization learning, thereby creating overregularization errors. U-shaped behaviour is indicative of this. In sum, the data we have for the errors of irregular verbs is partial evidence of a dual-mechanism. We still need information on the relationship between regularly inflected and irregularly inflected items to make a claim with any certainty.

There is also the matter of the contrast between the reaction times on accurate responses and overregularized responses. We saw how participants systematically reacted slower when producing overregularized responses in contrast to the accurate responses. This might be indirect evidence and support of the dual-mechanism account. Participants search the mental lexicon for an irregular past-tense form but cannot find one in some cases due to weaker memory traces. Consequently, they spend longer time giving a response to each task because they are trying to retrieve an irregular form from memory before applying the regular rule in the end.

Connectionism holds that the frequency of the regular pattern explains the 'defaultness' of the rule-like application of *-ed* for the past tense of verbs and *-s* to plural nouns. However, a German study by Marcus, Brinkmann, Clahsen, Wiese, & Pinker (1996) showed how the regular past tense formation in German (participle *-t*) applies to a smaller percentage of verbs than the English counterpart. Furthermore, it still acts as the default when encountering novel words. "The default suffixation is not due to numerous regular words reinforcing a pattern in associative memory" (Marcus et al., 1996, p. 190). Seeing as irregular items are more frequent than regular ones, the frequency of the regular pattern is not the reason that the regular inflection is more readily available as claimed by connectionists theorists. The previously mentioned study is a good example of this as opposed to English-based ones. Although this second research question is interesting from a theoretical perspective, I will have to refrain from exploring it any further due to the limitations of this study. It is nonetheless an interesting question to pursue in future research.

5.3 Limitations

Although we have suggestive evidence for an interval of a U-shaped curve, this interpretation must be met with caution. There are several factors which have likely contributed to the results of this study. Following are some considerations of the limitations and influencing factors that have likely impacted this study.

Due to the lack of available data from earlier age groups, the results in this study cannot firmly confirm nor disprove the existence of U-shaped learning in L2 acquisition (grammatical aspects). We might indeed be seeing an interval of the later stages of U-shaped learning, but more data is required to reveal what the developmental curvature looks like prior to the 8th grade. The grade-range can therefore be considered one of the limitations of this study. I discussed in 3.5. how the grade-range, 8th to 10th grade, might be too compact to see the entire U-shaped curve during L2 acquisition. This suspicion appears to be true. If we are indeed seeing the later stages of U-shaped learning, then this suggests that the onset of the curve begins earlier than in the 8th grade. This would at least be the case for this study's target group.

Moreover, the number of participants in each group is another factor which influences the results. A larger sample size makes it easier to make certain generalizations about the trends that we see in our random sampled data. Preferably, having 20 or more participants in all three groups would have been ideal. Also, having groups that are equal in size would have been preferable. A small number of samples might be the reason why the measurements are not normally distributed. As indicated by the Shapiro-Wilk test for accuracy and overregularization errors, many of the variables were not normally distributed. The Wilcoxon test was therefore preferable since it does not assume normal distributions. However, the Wilcoxon test has less power (Bayeen, 2008, p. 77). A larger sample size with normally distributed data would have permitted parametric tests with more power such as t-tests and the ANOVA-test.

Surprisingly, we did not find a significant difference between group 2 and 3 when comparing the overregularization rates for verbs in the easy condition of verbs, even though group 3 exhibited a floor effect (see figure 18). However, the medians are comparatively similar in both groups (median = 0 in both group 2 and 3). It is therefore not surprising that the p-value is rather high (p-value = 0.1198). This indicates that the two distributions overlap and are not significantly different. When performing the same test with the same variables for group 1 and 3, we saw what some might refer to as a borderline significant p-value (p-value = 0.05497), the rational being that it is close to being significant, but still not under the threshold of 0.05. The same is arguably the case when comparing the distributions between group 2 and 3 for the overall overregularization rate (p-value = 0.06649). The NULL-hypothesis, that the two sampled distributions are the same, is consequently close to being rejected in both cases. As stated, we are dealing with a rather small number of participants. It is therefore conceivable that a larger sample size might have given clearer results. Further studies are therefore needed to investigate this more in-depth.

Although we do have supportive and indicative evidence for U-shape learning with the use of grade/proficiency as a proxy to age, we also saw a lot of variability in many of the measurements and variables within each group. Separating and organizing the groups

according to levels of proficiency within the age ranges might have yielded an even clearer picture of the developmental curvature that we are interested in. Since L2 acquisition does not develop as uniformly as in L1 acquisition, sorting, matching, and grouping the participants according to proficiency might have been advantageous. However, doing so would have required testing and assessing proficiency levels prior to administering the elicitation tasks which goes beyond the boundaries of the economical limits and the timeframe of this thesis. Moreover, sorting participants according to proficiency level is not without problems either. There is a trade-off since we lose some of the benefits of having a random sampled dataset. Random sampled data enable us to be more confident that our results are representative of the entire population. Matching participants adds the risk of creating a biased sampling. Regardless, it is still striking that we see a floor effect in group 3 for overregularization rates of easy verbs and clear significant differences between the groups for accuracy rates.

6 Conclusion

This study set out to explore whether Norwegian L2 learners of English exhibit a similar learning curve in the acquisition of irregular and regular aspects of verb morphology as L1 learners of English do. We have suggestive evidence that our group of Norwegian L2 learners of English do in fact display later stages of U-shaped learning behaviour in the acquisition of verbs. Accuracy rates increased with grade, whereas overregularization rates decreased. Furthermore, group 3 exhibited a floor effect for the overregularization rate of highly frequent verbs. A Wilcoxon test looking at the overlap between the distributions found a borderline significant difference between group 1 and 3 for the rate of overregularization of highly frequent verb items. This result can likely be ascribed to the small sample size, and we speculate that we could have found a smaller p-value marking a clear significant difference by having a larger test group. Our results speak in favour of the universality of U-shaped learning. However, further studies are indeed required to confirm this.

The second aim of this study was to investigate if we found any evidence for a dual-organization of the L2 lexicon. The findings in this study suggest that a dual-organization of the mental lexicon is more plausible than a connectionist account. This is due to the patterns in overregularization errors which are better explained by a process of restructuring and acquisition of the regular past tense rule. On the other hand, we cannot rule out the possibility of a connectionist model seeing as we do not have any data on how the participants behave when producing regularly inflected verbs. Studies demonstrating a disassociation between lexical memory and productive grammar in the L2, similar to Rodriguez-Fornells et al. (2001) and Clahsen's (1999) ERP-studies, could give us a more definite answer to this question.

Our results speak in favour of the universality of U-shaped learning. However, further studies are indeed required to confirm this. I therefore suggest that a within-participant longitudinal study, similar to Geeslin & Guijarro-Fuentes' (2006) study but pertaining to grammar acquisition, would be beneficial to explore this phenomenon in further detail. This current study should also be repeated including data from earlier grades to see if we can observe the entire U-shaped curve. Additionally, testing English L2 learners but with a typologically dissimilar L1 such as Chinese Mandarin would be valuable to establish the universality of the phenomenon. The reason being that Chinese Mandarin, unlike Norwegian, does not share a similar inflectional paradigm of irregular and regular verbs. This would rule out the possibility of L1-transfer.

References

Bateman, W. G. (1916). The language status of three children at the same ages. *Pedagogical Seminary*. 23, 211-240.

Bayeen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge, UK: Cambridge University Press.

Behrens, H., & Tomasello, M. (1999). And what about the Chinese? *Behavioral and Brain Sciences*, 22, 1014.

Berko, J. (1958). The Child's Learning of English Morphology. *WORD*, 14(2-3), 150-177.

Bybee, J. L., & Moder, C. L. (1983). Morphological Classes as Natural Categories. *Language*, 59 (2), 251-270.

Cazden, C. B. (1968). The Acquisition of Noun and Verb Inflections. *Child Development*, 39(2), 433-448.

Carlucci, L., & Case, J. (2013). On the Necessity of U-Shaped Learning. *Topics in Cognitive Sciences*, 5, 56-88.

Clahsen, H. (1999). Lexical entries and rule of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences*, 22, 991-1013.

Ellis, N. C. (2003). Constructions, Chunking, and Connectionism: The Emergence of Second Language Structure. In C. J. Doughty & M. H. Long (Ed.), *The Handbook of Second Language Acquisition*. UK: Blackwell Publishing Ltd..

Ervin, S. M., & Miller, W. R. (1963). Language development. In H. W. Stevenson (Ed.), *Child psychology: The sixty-second yearbook of the National Society for the Study of Education* (Pt. 1). Chicago: University of Chicago Press.

Evensen, M. W. (2014). *Searching for Patterns in English as a Second Language: A Norwegian study of 10 and 15 year old learners* (master's thesis). Norwegian University of Science and Technology, Trondheim.

Gass, S. M., & Selinker, L. (2008). *Second Language Acquisition. An introductory course* (3rd ed.). London, UK: Routledge.

Geeslin, K. L., & Guijarro-Fuentes, P. (2006). *A Longitudinal Study of Copula Choice: Following development in variable structures*. Unpublished manuscript. Indiana University, University of Plymouth.

Grosjean, F. (1992). Another View of Bilingualism. In R. J. Harris (Ed.), *Cognitive Processing in Bilinguals*. Netherlands: Elsevier Science Publishing Company Inc.

Grønning, M. M. (2016). *An investigation of Norwegians' L2 knowledge of English verb Phrase Anaphora Patterns* (master's thesis). Norwegian University of Science and Technology, Trondheim.

Jackendoff, R. (2009). *Foundations of Language. Brain, Meaning, Grammar, Evolution*. New York: Oxford University Press.

Jacobson, P. F., & Yu, Y. H. (2018). Changes in English Past Tense Use by Bilingual School-Age Children With and Without Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research*, 61(10), 2532-2546.

Jensen, I. N. (2016). *The Bottleneck Hypothesis in L2 acquisition: Norwegian L1 speakers' knowledge of syntax and morphology in English L2* (master's thesis). UiT The Artic University of Norway, Tromsø.

Karmiloff, K., & Karmiloff-Smith, A. (2002). *Pathways to Language: From Fetus to Adolescent*. United States of America: Harvard University Press.

Lardiere, D. (2003). Second Language Knowledge of [+/-Past] vs. [+/-Finite]. *Proceedings of the 6th Generative Approach to Second Language Acquisition Conference (GASLA 2002)*, 176-189.

Long, M. (1990). Maturational Constraints on Language Development. *Studies in Second Language Acquisition,* 12(3), 251-285.

Marcovitch, S., & Lewkowicz, D. J. (2004). U-Shaped Functions: Artifact or Hallmark of Development? *Journal of Cognition and Development*, 5(1), 113-118.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., Xu, F., & Clahsen, H. (1992). Overregularization in Language Acquisition. *Monographs of the Society for Research in Child Development*, 57(4), i+iii+v-vi+1-178.

Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1996). German Inflection: The exception that proves the rule. *Cognitive Psychology*, 29, 189-256.

McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections: what does the evidence rule out? *TRENDS in Cognitive Sciences*, 6(11), 465-472.

Pinker, S. (1991). Rules of Language. *Science*, 253, 530-535.

Pinker, S. (1998). Words and Rules. *Lingua*, 106, 219-242.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73-193.

Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *TRENDS in Cognitive Sciences*, 6(11), 456-463.

Plunkett, K., & Marchman, V. (1991). U-shaped learning and frequency effects in a mulit-layered perceptron: Implications for child language acquisition. *Cognition*, 38, 43-102.

Rodriguez-Fornells, A., Clahsen, H., Lleó, C., Zaake, W., & Münte, T. F. (2001). Event-related brain responses to morphological violations in Catalan. *Cognitive Brain Research*, 11, 47-58.

Rumelhart, D., & McClelland, J. L. (1986). On learning the past tense of English verbs. Implicit rules or parallel distributed processing? In J. L. McClelland, D. Rumelhart & the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition*. MIT Press: Cambridge, MA.

Shirai, Y. (1990). *U-shaped Behavior in L2 Acquisition*. Unpublished manuscript. University of California, Los Angeles.

Slabakova, R. (2016). *Second Language Acquisition*. England: Oxford University Press.

Sonnenstuhl, I., & Huth, A. (2002). Processing and Representation of German *-n* Plurals: A Dual Mechanism Approach. *Brain and Language*, 81, 276-290.

Strætkvern, S. B. (2017). *Left Dislocation in L2 English. L1-effects at the syntax-discourse interface* (master's thesis). Norwegian University of Science and Technology, Trondheim.

Utdanningsdirektoratet. (2013). *Læreplan i engelsk* (ENG1-03). Retrieved September 17, 2019 from https://www.udir.no/kl06/ENG1-03.

Vulchanova, M., Talcott, J. B., Vulchanov, V., Stankova, M., & Eshius, H. (2013). Morphology in autism spectrum disorders: Local processing bias and language. *Cognitive Neuropsychology*, 29(7-8), 585-600.

Williams, S. (2018). *Exam SPRÅK3300 – H18*. Unpublished paper. Norwegian University of Science and Technology, Trondheim.

APPENDIX A

Information and consent sheet for students in 8th and 9th grade

Kjære deltager

Jeg er en lektorstudent som skriver min master innen tilegnelse av språk. Som beskrevet i infoskrivet som dine foresatte har skrevet under så er jeg interessert i å vite mer om din språklige utvikling.

Testen vil bestå av tre deler. Den første delen er relatert til deg, din språklige bakgrunn, og vaner rundt din bruk av engelsk. Den andre delen vil være oppgaver hvor du vil bli bedt om å fylle inn ord du anser som passende for setningen. Tilslutt, så vil du få lignende oppgaver som i del to.

Selv om foreldrene dine har skrevet under for at de godkjenner at du deltar i denne testen, så er det frivillig hvorvidt du selv vil bli med. Begynner du testen men finner ut at du ønsker å trekke deg, er det bare å avbryte og opplyse meg (Stefan Williams) at du ikke ønsker å delta.

I sammenheng med denne testen innhenter vi opplysninger om dine foreldre og deg. Vi spør også dine foreldre om opplysninger om deg. Nedenfor ser du en liste med type opplysninger vi innhenter i prosjektets helhet:

- Navn (Vil være adskilt fra data og ikke lagret sammen)
- Kjønn
- Fødselsdato
- Språklig bakgrunn
- Helseopplysninger som kan være relevante for språk
- Foreldres utdanning og yrker

Har du noen spørsmål er det bare å ta kontakt!

Med vennlig hilsen,

Stefan Williams

Mastergradsstudent i lektorprogrammet
NTNU, Institutt for Språk og Litteratur
Tlf. 97426755, e-post: stefandw@stud.ntnu.no

SAMTYKKEERKLÆRING

☐ Ja, jeg samtykker å delta i dette forskningsprosjektet.

APPENDIX B

Information and consent sheet for parents

Kjære foresatte

Jeg er en lektorstudent ved NTNU som skriver min masteroppgave om språkutvikling innen tilegnelse av engelsk som andrespråk. I den anledning trenger jeg testpersoner med norsk som et av sine morsmål og engelsk som et andrespråk fra skolegangen sin. Av den grunn ønsker jeg å be om tillatelse å teste deres barn to ganger i løpet av det kommende året.

Med «test» menes ikke det samme som «å prøve», men heller i en forskningsmessig forstand. Jeg er ikke interessert i å bedømme riktige eller gale svar her, men jeg vil heller ha en mer beskrivende tilnærming til svarene deres. Det jeg er interessert i er å få statistisk data på deres språkutvikling innen engelsk. Jeg vil benytte med av en velutprøvd testmetode som er svært overkommelig for ungenes del. Testen vil være digitalisert gjennom nettplattformen SurveyGizmo.

Hver test vil ta ca. en halv time per gang. Videre så ønsker jeg å gjennomføre den samme testen to ganger for å ha nok sammenligningsmateriale. Av den grunn er det ønskelig å gjennomføre testen én gang i slutten av september, og én gang til rundt desember. Helhetlig sett burde dette ikke ta mer enn en times tid.

Prosjektet er designet av meg i dialog med min veileder professor Mila Vulchanova ved NTNU. Selve gjennomføringen av testene vil bli utført av meg. I tillegg er prosjektet meldt til NSD.

Som nevnt er jeg ute etter statistiske data, ikke etter å se hvor "flinke" de enkelte barna er. Resultatene av testingen vil bare håndteres av meg og min veileder, og vil være fullstendig anonymiserte før de brukes i offentligheten. I første omgang lagres alle resultatene med en personkode som tilsvarer hver elev. Personkoden vil være før i en separat navneliste slik at navn på eleven og resultater på testene ikke oppbevares på samme sted. Navnelisten med personkodene vil være lagret i fysisk form, mens resultatene vil være lagret digitalt. Ved prosjektets slutt i desember 2019 vil dataene anonymiseres fullstendig ved at elevenes navn og andre personopplysninger slettes helt.

Deltakelse i prosjektet er selvsagt frivillig, og hvis dere ikke svarer på denne forespørselen vil deres barn ikke bli involvert i studien. Hvis dere har spørsmål eller ønsker mer informasjon om prosjektet, må dere gjerne kontakte meg. Godtar dere at barnet deres deltar i prosjektet, kan dere selvsagt likevel ombestemme dere og reservere dere når som helst. Dere behøver i så fall ikke gi noen begrunnelse for hvorfor dere ønsker å trekke dere.

Hvis dere velger alternativet "ja" under, vil dere i løpet av kort tid få et enkelt spørreskjema der jeg ber om litt mer informasjon om barnets bakgrunn som kan være relevant for språklæringen. Vi gjør oppmerksom på at vi etterspør helseopplysninger som er relevante for barnets språkutvikling. Denne informasjonen vil behandles konfidensielt på lik linje med all annen personlig informasjon som kommer fram gjennom prosjektet. Kun jeg, veileder, og en intern medarbeider fra instituttet vil ha tilgang til informasjonen gitt i sammenheng med prosjektet.

Dine rettigheter

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

- Innsyn i hvilke personopplysninger som er registrert om deg
- Å få rettet personopplysninger om deg
- Å få slettet personopplysninger om deg
- Få utlevert en kopi av dine personopplysninger (dataportabilitet)

- Å sende klage til personvernombudet eller Datatilsynet om behandlingen av dine opplysninger.

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

Ønsker du å kontakte veileder?
Hvis dette er tilfellet, vennligst ta kontakt med

Mila Dimitrova Vulchanova
Professor ved Institutt for Språk og Litteratur
e-post: mila.vulchanova@ntnu.no
Tlf: 73596791
Bygg 5, 5555, Dragvoll

Ønsker du å kontakte NTNU sitt personvernombud?
Hvis dette er tilfellet, vennligst ta kontakt med

Thomas Helgesen
Personvernombud ved NTNU
Tlf: 93079038
e-post: thomas.helgesen@ntnu.no
Adresse: Sluppenveien 12B/C, Møllenberg 4 etg, Trondheim

Med vennlig hilsen,

Stefan Williams

Mastergradsstudent i lektorprogrammet
NTNU, Institutt for Språk og Litteratur
Tlf. 97426755, e-post: stefandw@stud.ntnu.no

SAMTYKKEERKLÆRING

☐ Ja, jeg samtykker i at mitt barn deltar i forskningsprosjektet og det følgende:

- ☐ At mitt barn svarer på et elektronisk spørreskjema
- ☐ At jeg svarer på et papirbasert spørreskjema
- ☐ At det blir hentet inn opplysninger om meg som tredjeperson

_____	_____	_____
Barnets navn	Sted og dato	Foresatt 1 underskrift

_____
Foresatt 2 underskrift

Opplysninger om barnet og barnets språkbakgrunn/-ferdighet

Hvis du godtar at barnet deltar i prosjektet, krysser du av for de alternativene nedenfor som passer best for ditt barn:

Barnets fødselsdato: _____

☐ Barnet har deltatt i vanlig engelskundervisning siden skolestart. Hvis nei, beskriv under:

☐ Barnet har kun norsk som sitt morsmål. Hvis nei, eller at barnet har flere morsmål, beskriv hvilke eventuelle språk under:

☐ Barnet har ingen hørsels-/språklige/motoriske problemer. Hvis ja, beskriv gjerne under:

_____ _____ _____

Barnets navn Sted og dato Foresattes underskrift

Bruk gjerne kommentarfeltet under til å oppgi flere opplysninger, f.eks. bakgrunnen for at barnet evt. er tospråklig og om nivået på språkene, eller f. eks. hvis barnet har hatt lengre opphold i engelskspråklige land eller lignende.

Opplysninger om foresattes bakgrunn

Hvis du godtar at barnet deltar i prosjektet så setter vi stor pris på om du kan oppgi noen opplysninger om foresattes bakgrunn.

Begge foresattes utdanning:

Begge foresattes yrker:

Barnets navn	Sted og dato	Foresattes underskrift

APPENDIX C

Information and consent sheet for 16-year-old participants

Kjære deltager

Jeg er en lektorstudent ved NTNU som skriver min masteroppgave om språkutvikling innen tilegnelse av engelsk som andrespråk. I den anledning trenger jeg testpersoner med norsk som et av sine morsmål og engelsk som et andrespråk fra skolegangen sin. Av den grunn ønsker jeg å invitere deg til å delta i testen min.

Med «test» menes ikke det samme som «å prøve», men heller i en forskningsmessig forstand. Jeg er ikke interessert i å bedømme riktige eller gale svar her, men jeg vil heller ha en mer beskrivende tilnærming til svarene. Det jeg er interessert i er å få statistisk data på din språkutvikling innen engelsk. Jeg vil benytte med av en velutprøvd testmetode som er svært overkommelig for testpersonenes del. Testen vil være digitalisert gjennom nettplattformen SurveyGizmo.

Hver test vil ta ca. en halv time per gang. Videre så ønsker jeg å gjennomføre den samme testen to ganger for å ha nok sammenligningsmateriale. Helhetlig sett burde dette ikke ta mer enn en times tid.

Prosjektet er designet av meg i dialog med min veileder professor Mila Vulchanova ved NTNU. Selve gjennomføringen av testene vil bli utført av meg. I tillegg er prosjektet meldt til NSD.

Som nevnt er jeg ute etter statistiske data, ikke etter å se hvor "flinke" deltagerne er. Resultatene av testingen vil bare håndteres av meg og min veileder, og vil være fullstendig anonymiserte før de brukes i offentligheten. I første omgang lagres alle resultatene med en personkode som tilsvarer hver elev. Personkoden vil være ført i en separat navneliste slik at navn på deltagerne og resultater på testene ikke oppbevares på samme sted. Navnelisten med personkodene vil være lagret i fysisk form, mens resultatene vil være lagret digitalt. Ved prosjektets slutt i desember 2019 vil dataene anonymiseres fullstendig ved at deltagernes navn og andre personopplysninger slettes helt.

Deltakelse i prosjektet er selvsagt frivillig. Hvis du har spørsmål eller ønsker mer informasjon om prosjektet, må du gjerne kontakte meg. Godtar du å delta i prosjektet, kan du selvsagt likevel ombestemme deg og reservere deg når som helst. Du behøver i så fall ikke gi noen begrunnelse for hvorfor du ønsker å trekke deg.

Hvis du velger alternativet "ja" under, vil du i løpet av kort tid få et enkelt spørreskjema der jeg ber om litt mer informasjon om din bakgrunn som kan være relevant for språklæring. Kun jeg, veileder, og en intern medarbeider fra instituttet vil ha tilgang til informasjonen gitt i sammenheng med prosjektet.

Dine rettigheter

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

- Innsyn i hvilke personopplysninger som er registrert om deg
- Å få rettet personopplysninger om deg
- Å få slettet personopplysninger om deg
- Få utlevert en kopi av dine personopplysninger (dataportabilitet)
- Å sende klage til personvernombudet eller Datatilsynet om behandlingen av dine opplysninger.

Hva gir oss rett til å behandle dine opplysninger?

Vi behandler opplysninger om deg basert på ditt samtykke.

APPENDIX D

Full survey for participants in 8th and 9th grade

Masteroppgave - Språkkunnskap

Part 1: Background and participant information

1. Please submit your participant ID for this survey (For instance: ID1498): *

2. When were you born? *

3. Please specify your gender *

- ○ Male
- ○ Female

4. What is your native language (morsmål)? *

- ☐ Norwegian
- ☐ English
- ☐ Other

Please specify your other native language(s):

5. What other language(s) do you speak in addition to your native language(s)? Please specify language(s) and competence (high – medium – low) in the textbox below.

If you do not speak any additional languages, please write 'none'.

Examples of how to answer: English (medium) Arabic (high) Mandarin/Chinese (low) *

6. Do you have a diagnosis that could potentially affect your language learning (e.g. dyslexia, impaired hearing, etc.)? If yes, please specify. If no, simply write: 'no' *

7. Have you lived in an English speaking country for more than 2 months? If so, where and for how long? If no, simply write: 'no'. *

8. Do you have close family and/or friends that speak English? *

9. Approximately how many hours do you use English throughout a day? (Including at school, and other activities such as reading English online media etc.) *

10. When do you use English in everyday life? *

- [ ] While watching TV/movies/series with subtitles
- [ ] While watching TV/movies/series without subtitles
- [ ] When talking to friends/family
- [ ] At school
- [ ] When reading/writing English blogs/websites
- [ ] While reading English books (not including what you read at school)
- [ ] When writing English (not including assignments at school)
- [ ] While Playing games (PC-games, X-box, Playstation etc.)

11. Are you involved in one or more activities on your free time that involves use of English? (E.g. theater/role playing, gaming, etc.) *

- [ ] Yes, please specify
- [ ] No

Please specify which activities *

12. How would you rate your own English competence? *

- ☐ I can understand English text well
- ☐ I can understand English speech well
- ☐ I can write English well

13. Please rate how well you know English on a scale from 1-5. 1 = Poor, 5 = Very good *

	1	2	3	4	5
English	○	○	○	○	○

Task 1

Task 1:
SET: Every year, this guy sets out on a new journey.
What did he do last year?
He _____out on a journey. *

Task 2

Task 2:
RUN: Joey likes to run 10km every day.
What did Joey do yesterday?
He_____ 10km. *

Task 3

Task 3:
DO: Joey does something stupid every week.
What did Joey most likely do last week?
He____something stupid. *

Task 4

Task 4:
MAKE: Betty makes good food when Joey comes over for dinner.
Yesterday, Joey came over for dinner. What did Betty do?
She_____ good food for Joey. *

Task 5

Task 5:
GET: Every year, Joey gets Betty something nice for her birthday.
This year, Joey _____ Betty a nice present. *

Task 6

Task 6:
SAY: Whenever Joey and Betty argue, they say mean things to each other.
Last time they argued, both _____mean things to each other. *

Task 7

Task 7:
GO: Paul likes to go to the arcade to play games.
Paul _____ last Friday to the arcade at 12 o'clock. *

Task 8

Task 8:
TAKE: Paul takes Betty's sister Lisa to go for a hike in the weekends.
Last weekend, Paul _____ Lisa for a hike around a lake for a change. *

Task 9

Task 9:
SEE: Paul loves to see movies on his spare time.
Last night, Paul _____ Mad Max: Fury Road at home. *

Task 10

Task 10:
KNOW: Since Paul has seen the movie before, he knows what will happen.
One could say that Paul _____ what would happen at the end of the movie. *

Task 11

Task 11:
FIND: Betty's friend Anne likes to find nice stones along the beach.
Once, walking along the beach, Anne _____ a green stone! *

Task 12

Task 12:
COME: Anna usually comes over to Joey's birthday parties.
This time, Anna _____ late to the party. *

Task 13

Task 13:
THINK: Joey thinks that being late is rude
During yesterday's party, when Anna was late, Joey _____ that Anna was rude *

Task 14

TASK 14:
GIVE: Anna usually does not give presents.
At the party though, Anna _____ Joey a gift! *

Task 15

Task 15:
KEEP: Joey likes to keep his presents hidden away from people after parties.
After yesterday's party, Joey _____ his presents hidden. *

Task 16

TASK 16:
SHOW: Betty likes to show pictures of Joey at parties.
At yesterday's party, Betty had _____ all her pictures of Joey! *

Task 17

TASK 17:
FEEL: Usually, Joey feels happy when Betty shows people pictures of him.
However, at the party, Joey _____ embarrassed by some of the pictures. *

Task 18

Task 18:
LEAVE: Most of the time, Lisa leaves early when going to parties.
At yesterday's party, Lisa _____ early. *

Task 19

Task 19:
TELL: Paul likes to tell jokes at parties.
At yesterday's party, Paul _____ a funny joke that made everybody laugh! *

Task 20

Task 20:
PUT: Betty likes to put berries in top of cakes.
Betty _____ berries on top of the cake she baked last year. *

Task 21

Task 21:
HIDE: During hide and seek, you should hide from the others who are playing the game.
Lisa _____ an hour ago, and the others have still not been able to find her! *

Task 22

Task 22:
SLEEP: Sleeping is important.
However, when you have _____ for 15 hours, that's too long! *

Task 23

Task 23:
SING: Betty likes to sing.
Betty _____ on TV during last year's American Idol! *

Task 24

Task 24:
BLOW: Joey thinks it is hard to blow out all the candles on the cake.
This year through, Joey _____out all the lights on the cake! *

Task 25

Task 25:
STEAL: Paul likes to steal things.
However, during the party, Paul _____ Lisa's heart. *

Task 26

Task 26:
HURT: Paul hates to get hurt when playing sport.
Last time Paul played American football, Paul _____ his leg. *

Task 27

Task 27:
WAKE: Joey wakes up 7 o'clock every morning
However, Joey ____up at 1 o'clock last Sunday! *

Task 28

Task 28:
SHUT: When Paul enters his home, he usually shuts the door when he goes inside.
Yesterday, when Paul came home, he _____ the door after himself. *

Task 29

Task 29:
SHAKE: Joey likes to shake his milkshakes before he drinks them
Yesterday, Joey _____ his milkshake more than usual! *

Task 30

Task 30:
SLIDE: It's easy to slide on an icy road and fall over.
Last year, Anna _____on the ice and broke her leg! *

Task 31

Task 31:
SWIM: Betty likes to swim in the weekends.
Last weekend, Betty _____ 5km in a pool! *

Task 32

Task 32:
TEAR: Every Monday, Paul has to tear down all the posters from the poster-wall.
Last Monday, Paul _____ down 100 posters! *

Task 33

Task 33:
FREEZE: Brian likes to prepare all his food a week in advance, and freeze it afterwards.
Last night, after preparing his food, Brian _____ his food. *

Task 34

Task 34:
SHINE: Brian likes it when the sun shines on him.
Yesterday, the sun _____ through the clouds on Brian. *

Task 35

TASK 35:
BITE: Brian is afraid of getting bitten by zombies.
During a nightmare, Brian got _____ by a zombie! *

Task 36

Task 36:
BET: Brian has a gambling problem and likes to bet way too much money on poker.
Yesterday, Brian _____ 10 000kr on a poker game and lost! *

Task 37

Task 37:
FLY: In his spare time, Brian likes to fly.
Yesterday at the party, Brian _____ through the window. *

Task 38

Task 38:
LEND: People usually lend Brian money after he loses a poker game.
Joey _____ Brian 10 000kr after his last poker game. *

Task 39

Task 39:
SWING: Joey likes to play baseball and swing the baseball bat.
Yesterday, Joey _____ the bat and hit a homerun! *

Task 40

Task 40:
CREEP: Kevin likes to creep around the forest.
Last year, Kevin _____ around the forest and got in trouble afterwards! *

Task 41

Task 41:
FORBID: Anna's father likes to forbid Anna from doing certain things that he thinks are bad for her.
Last year, Anna's father _____ her from playing video games. *

Task 42

Task 42:
CLING: When rock climbing, Brian clings tightly onto the rocks.
Yesterday, Brian _____ tightly to the rocks so he would not fall. *

Task 43

Task 43:
BIDE: Brian, when playing poker, likes to bide his time.
During last night's game, Brian was patient and _____ his time. *

Task 44

Task 44:
CHILD: Betty gave birth to a child in August.
A year later, Betty gave birth to two more _____. *

Task 45

Task 45:
MAN: A baseball team cannot consist of only one man.
A baseball team consist of at least 9 _____. *

Task 46

Task 46:
WOMAN: In group one, there is 1 woman.
In group two, there are 4 _____. *

Task 47

Task 47:
LIFE: A regular person has one life.
A cat, on the other hand, has 9 _____. *

Task 48

Task 48:
PERSON: In room one, there is 1 person.
In room two, there are 4 _____. *

Task 49

Task 49:
GOOSE: Today, I only saw 1 goose.
Yesterday, I saw a flock of _____. *

Task 50

Task 50:
OX: Farmer Bob has 1 ox.
Farmer Brendon has 5 _____. *

Task 51

Task 51:

TOOTH: You are unlucky if you only have 1 tooth.

You are lucky if you have all your _____. *

Task 52

Task 52:

SHEEP: Farmer Bob has only 1 sheep.

Farmer Brendon has a herd of _____. *

Task 53

Task 53:

FISH: It would be unusual for a stream to only have 1 fish.

Usually, there are many _____ in a stream! *

APPENDIX E

Figures

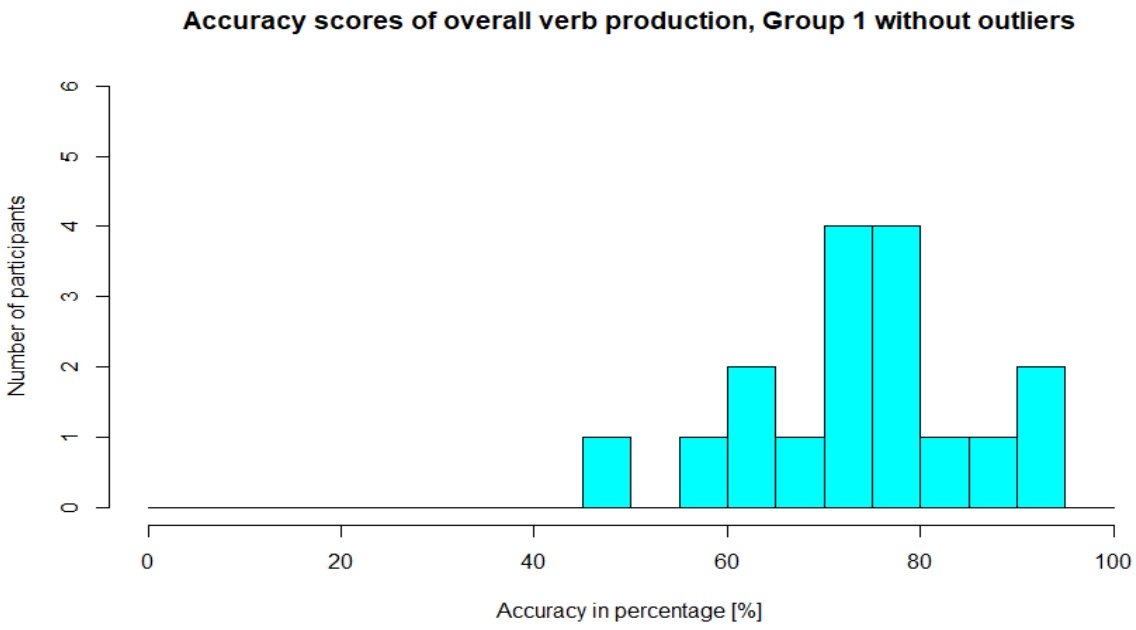


Figure 1: Overall accuracy of verbs in group 1.

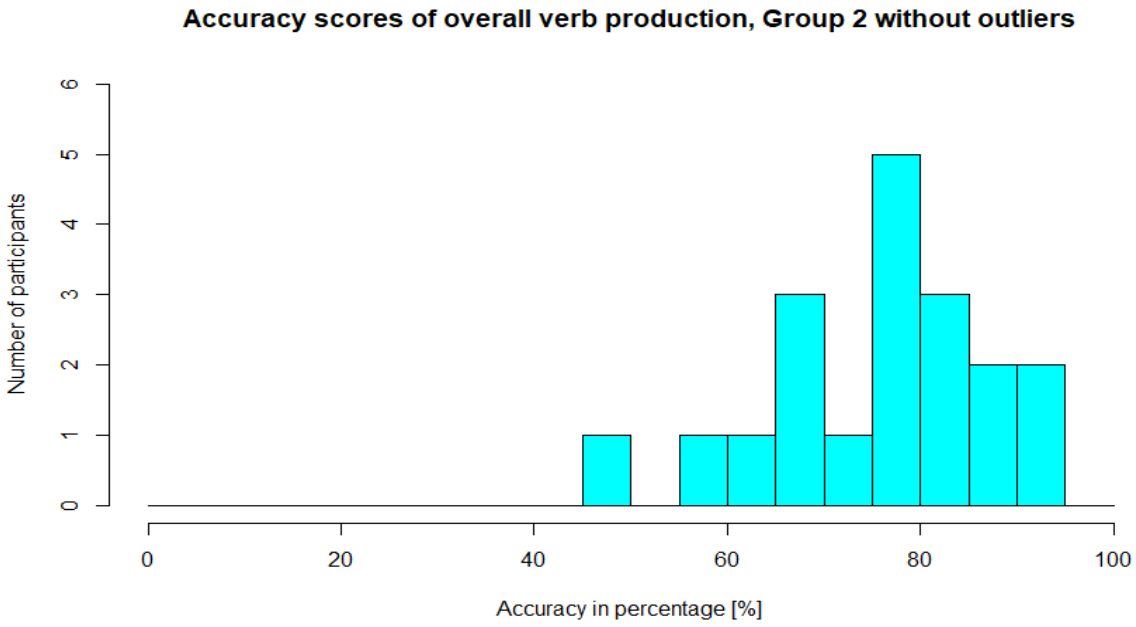


Figure 2: Overall accuracy of verbs in group 2.

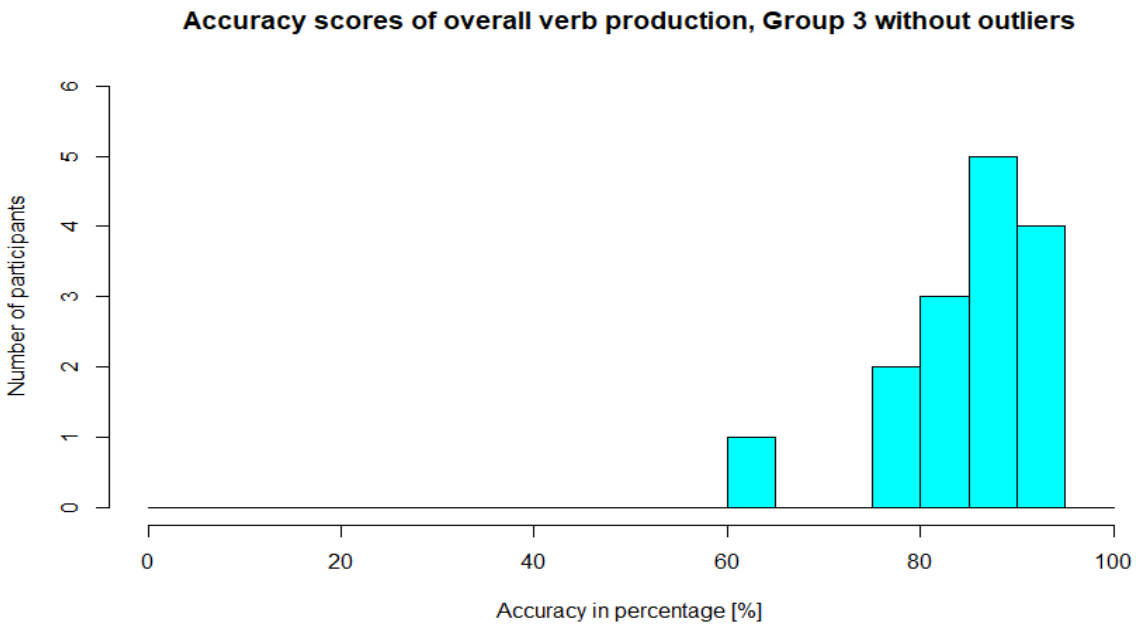


Figure 3: Overall accuracy of verbs in group 3.

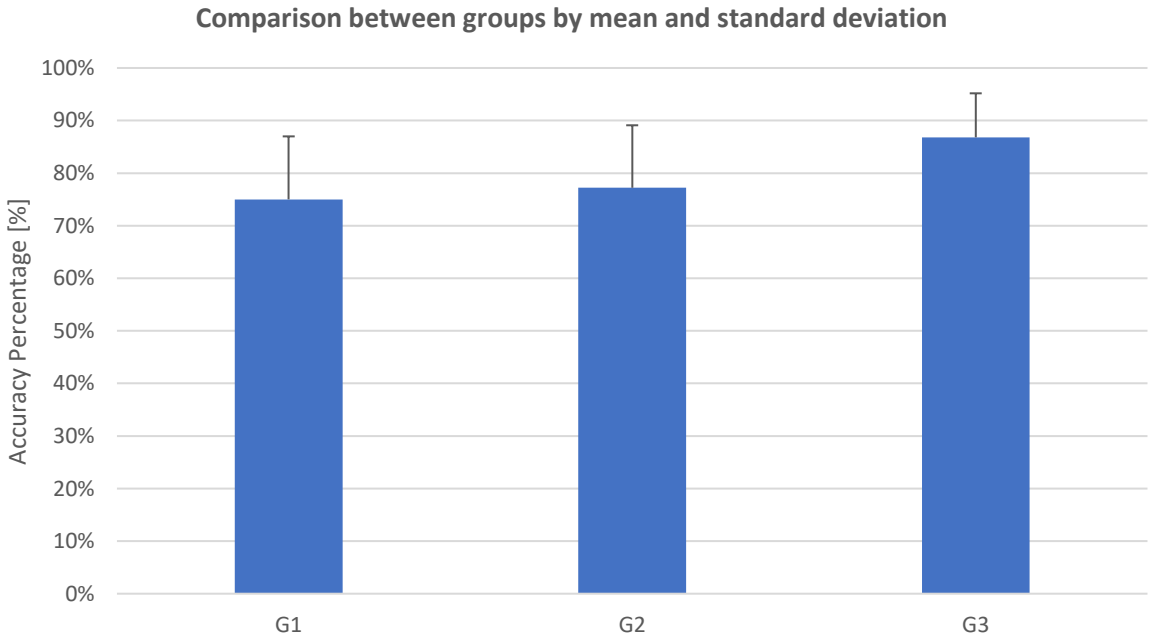


Figure 4: Comparison between groups by mean and standard deviation for overall accuracy of verbs.

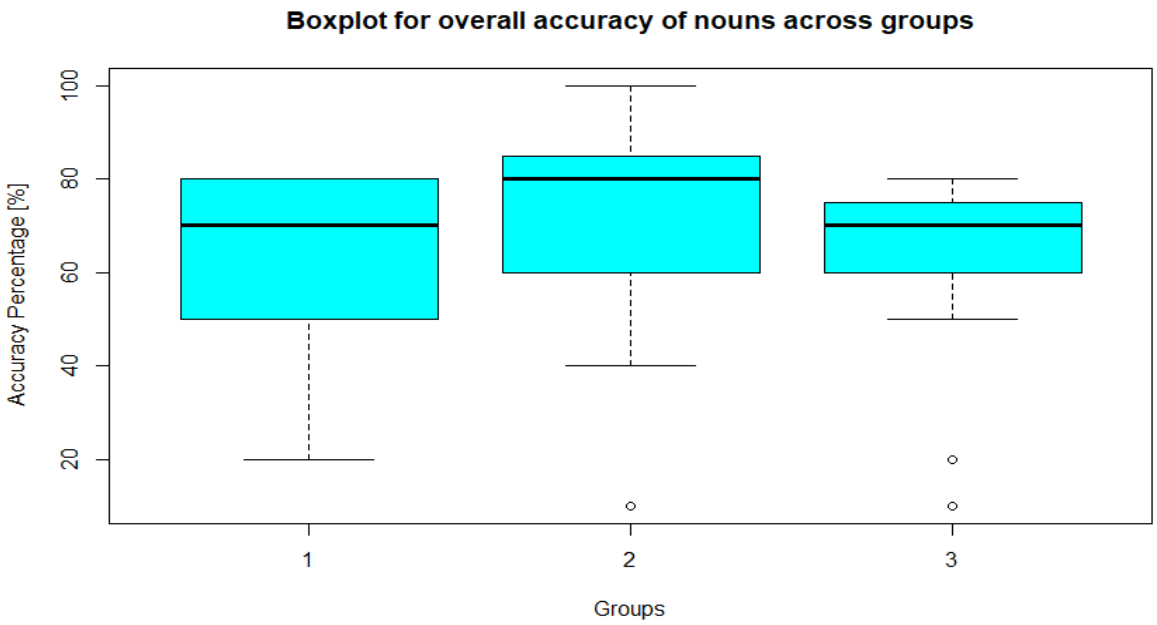


Figure 5: Comparison between groups for overall accuracy of nouns.

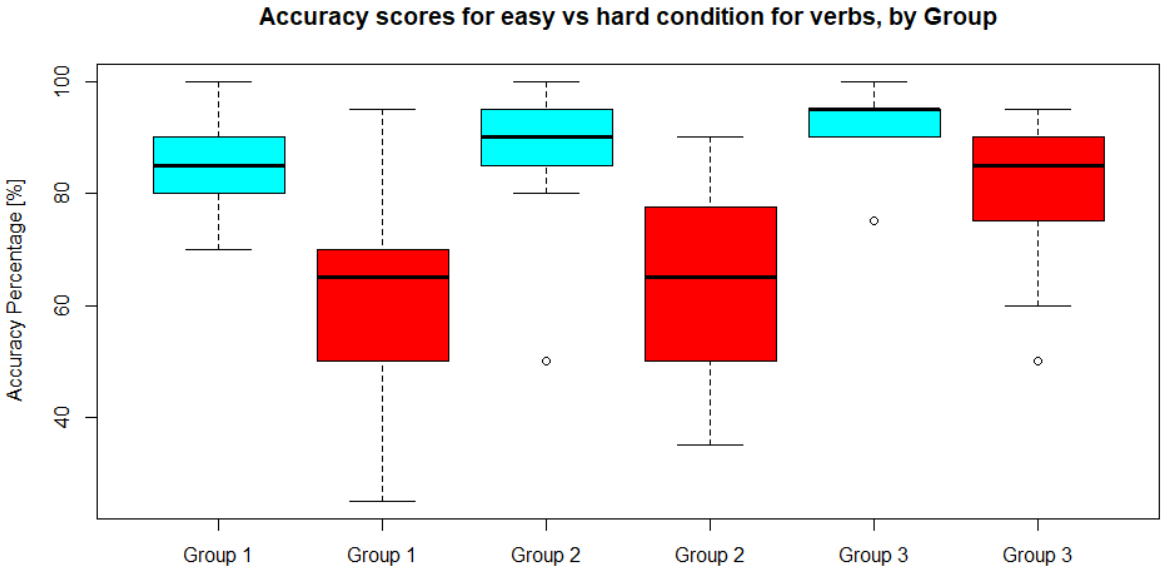


Figure 6: Accuracy scores for easy (turquoise) vs hard (red) condition for verbs, by Group.

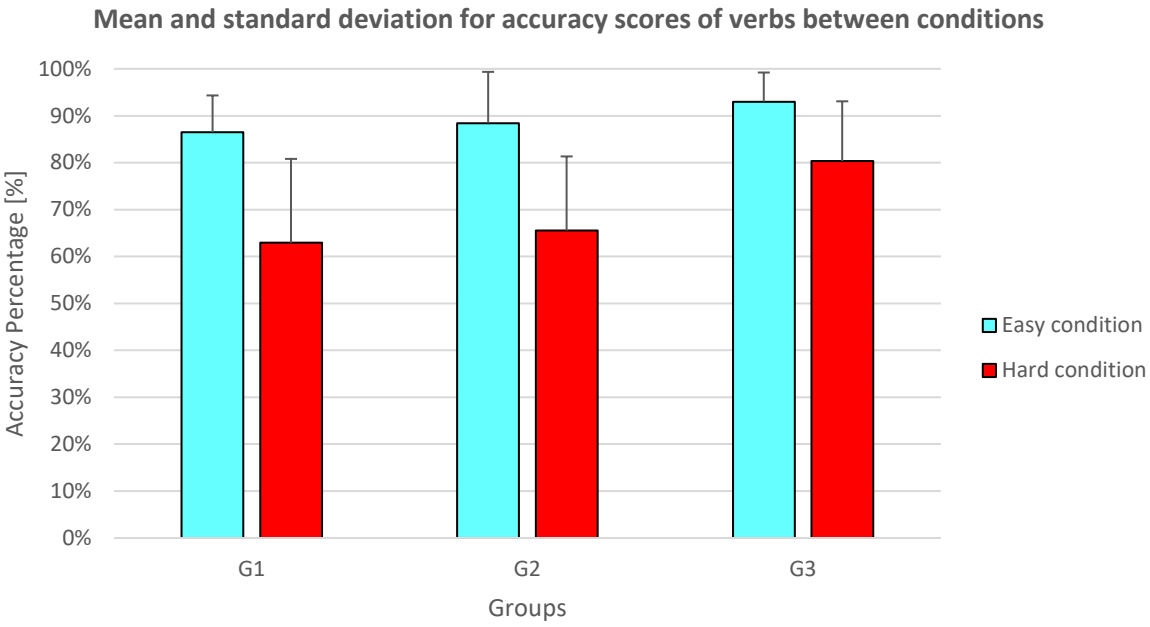


Figure 7: Comparison between groups for mean scores and standard deviation, easy and hard condition of verbs.

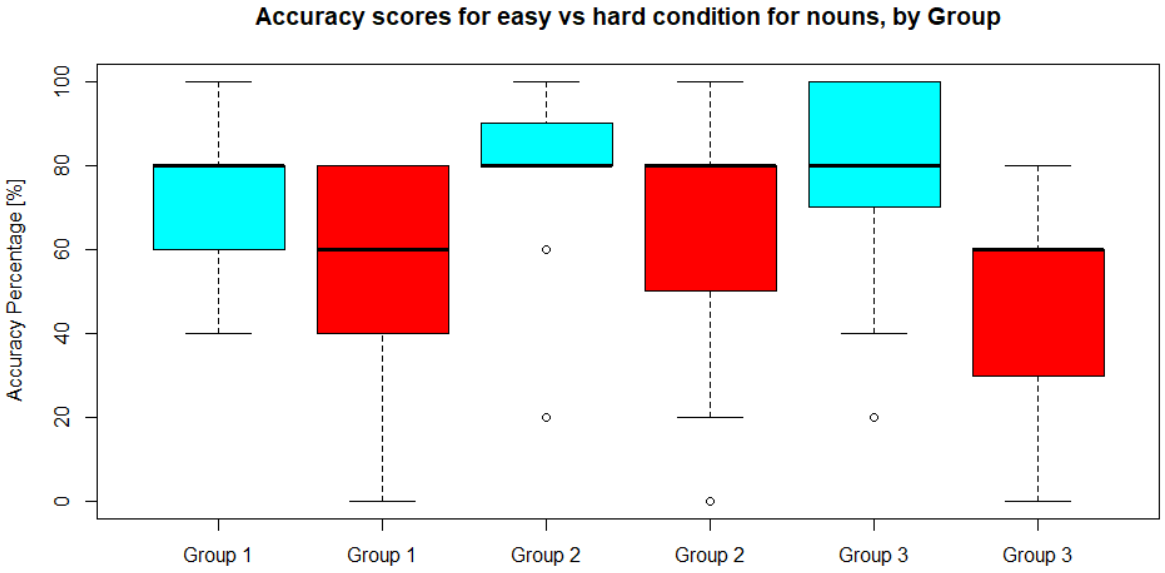


Figure 8: Accuracy scores for easy vs hard condition for nouns, by Group.

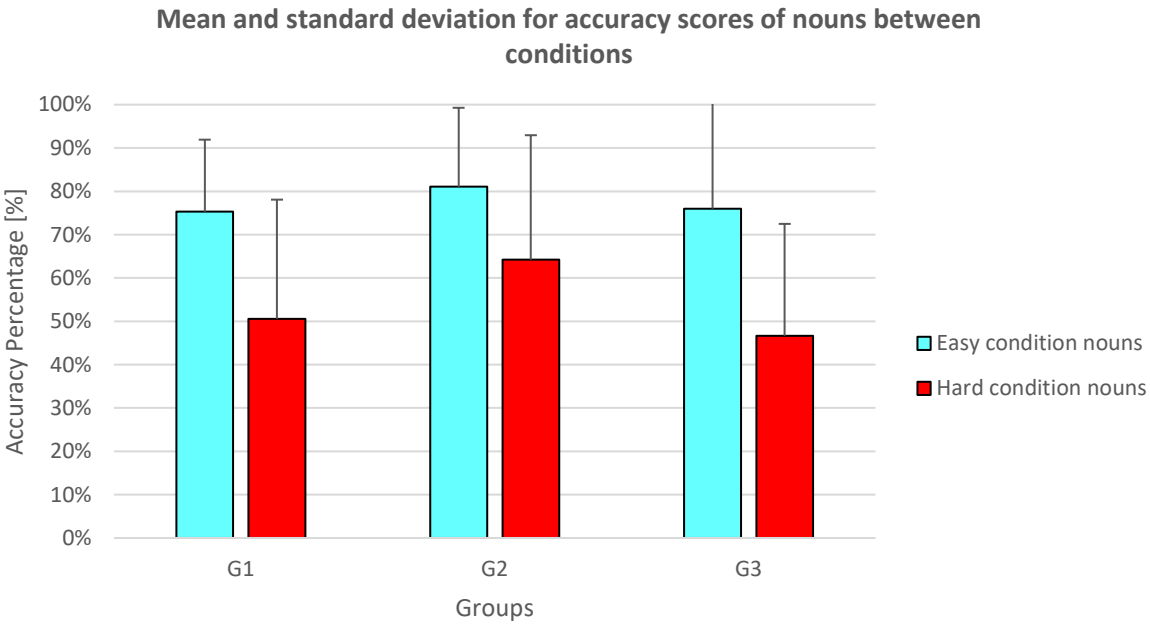


Figure 9: Comparison between groups for mean scores and standard deviation, easy and hard condition of nouns.

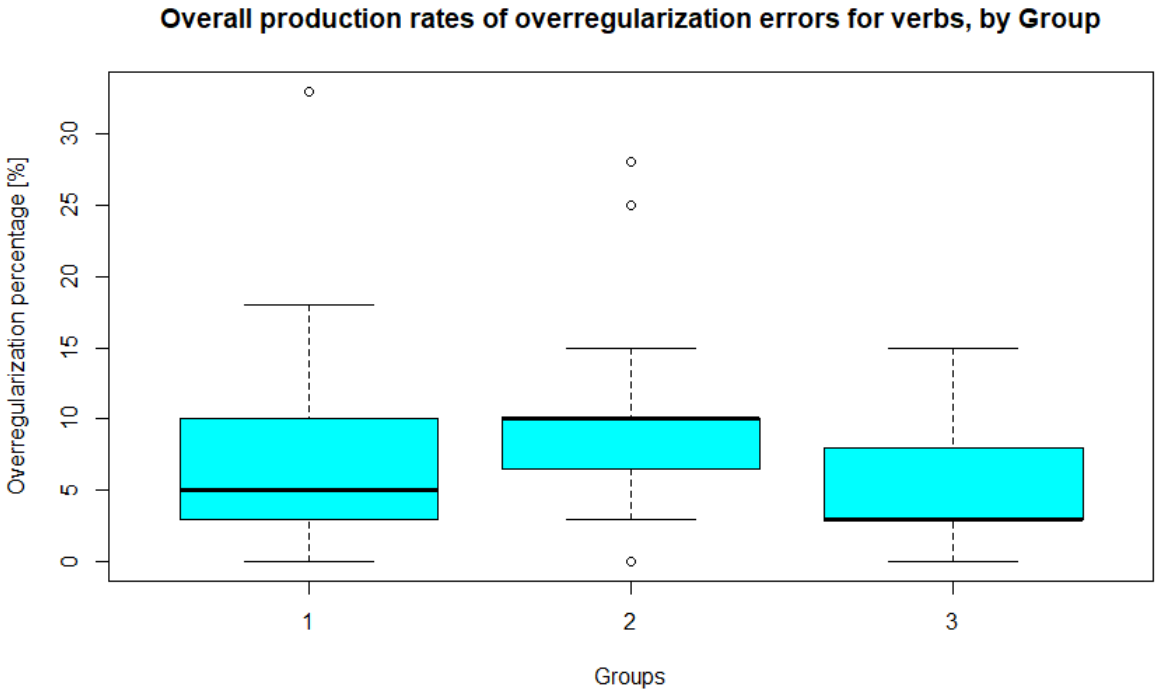


Figure 10: Overall production rates of overregularization errors for verbs, by Group.

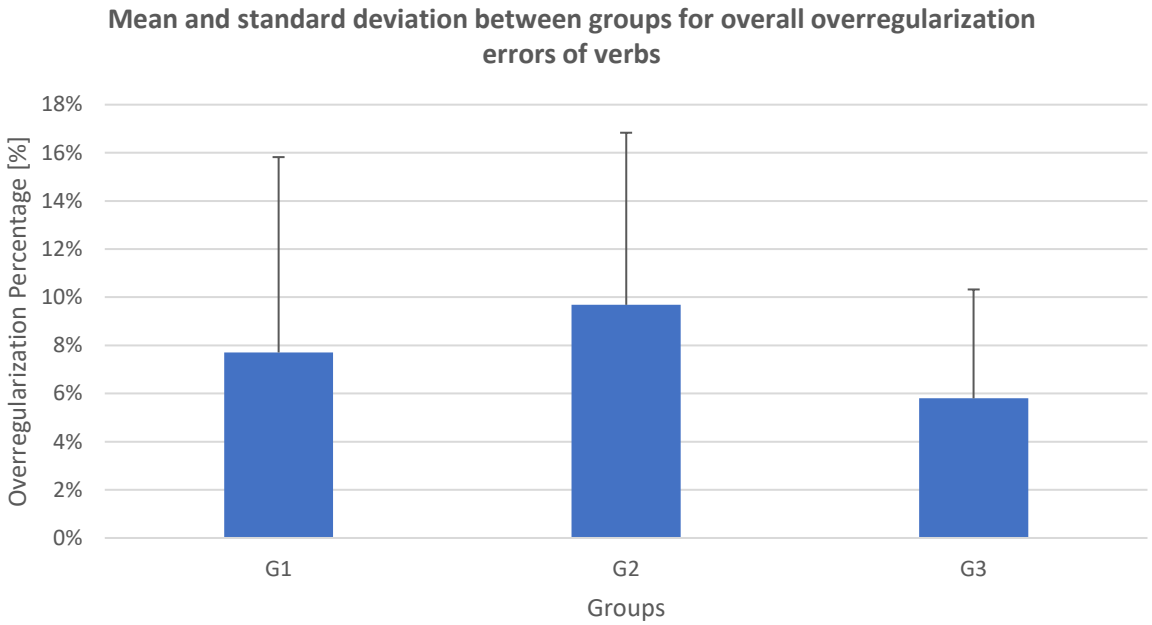


Figure 11: Comparison between groups for overall production rate of overregularization errors for verbs.

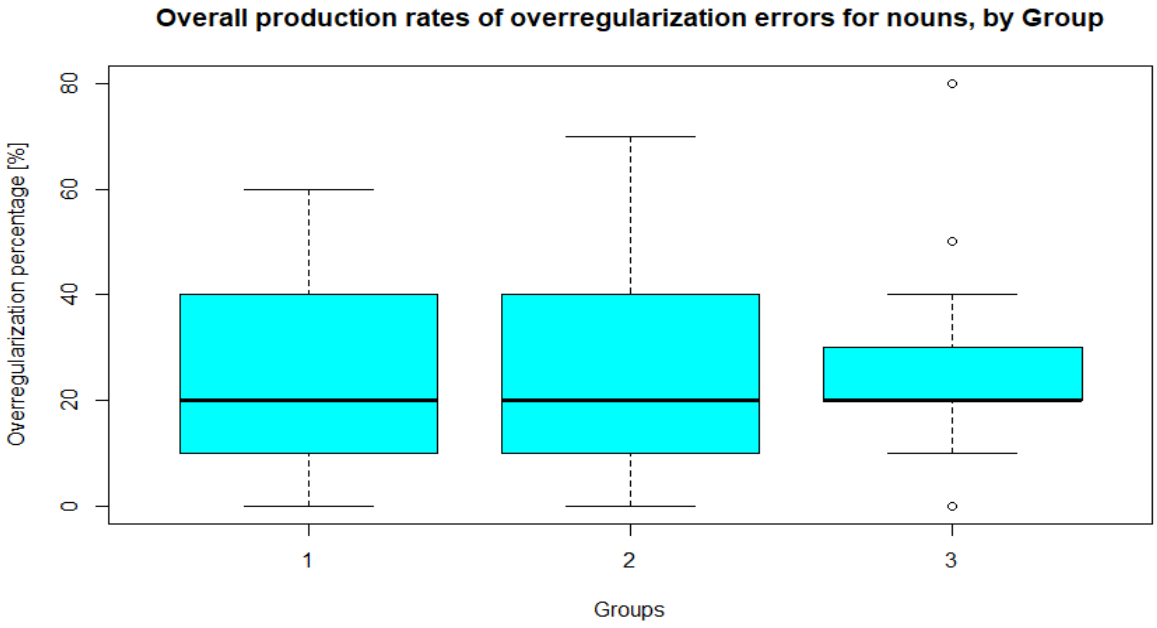


Figure 12: Overall production rates of overregularization errors for nouns, by Group.

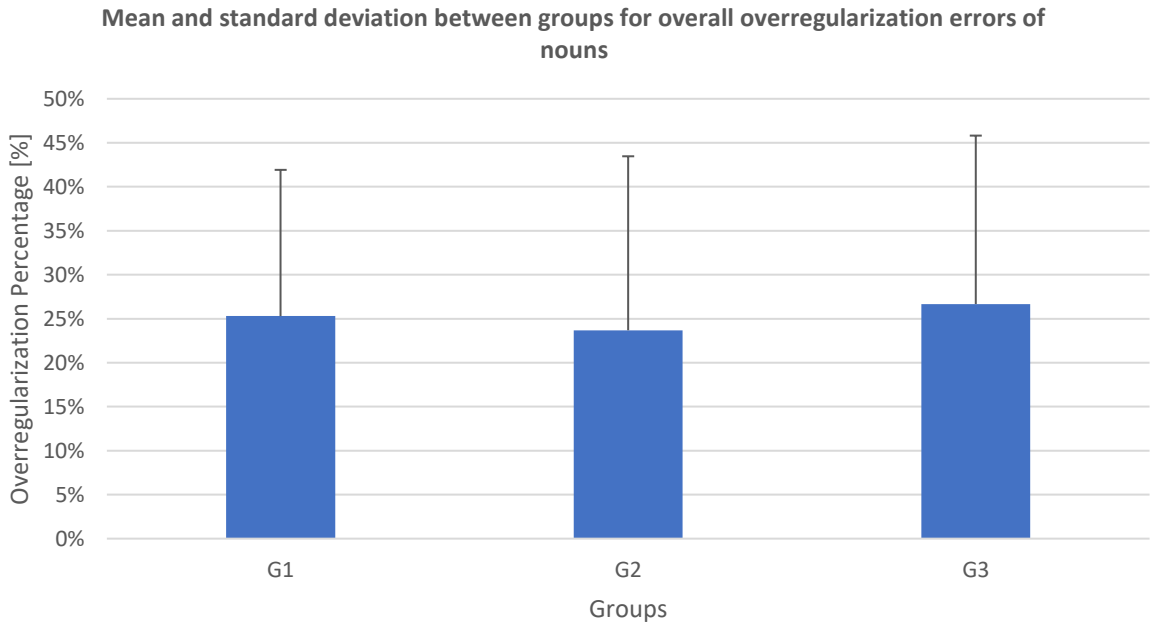


Figure 13: Comparison between groups for overregularization errors of nouns.

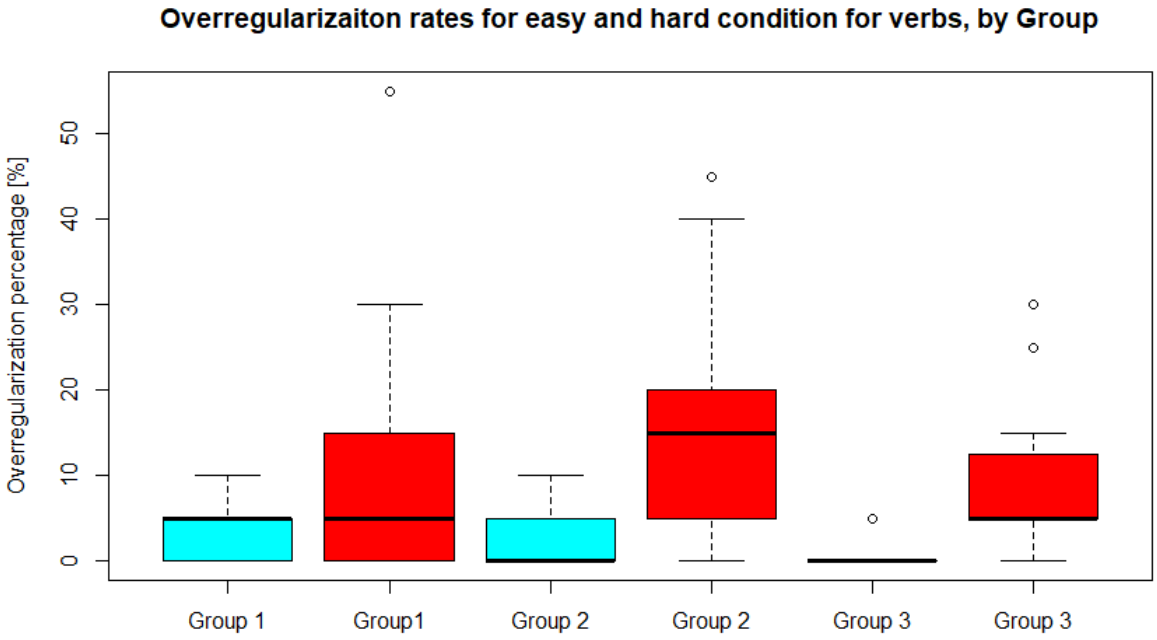


Figure 14: Comparison between easy condition (turquoise) and hard condition (red) for rate of overregularization of verbs, by Group.

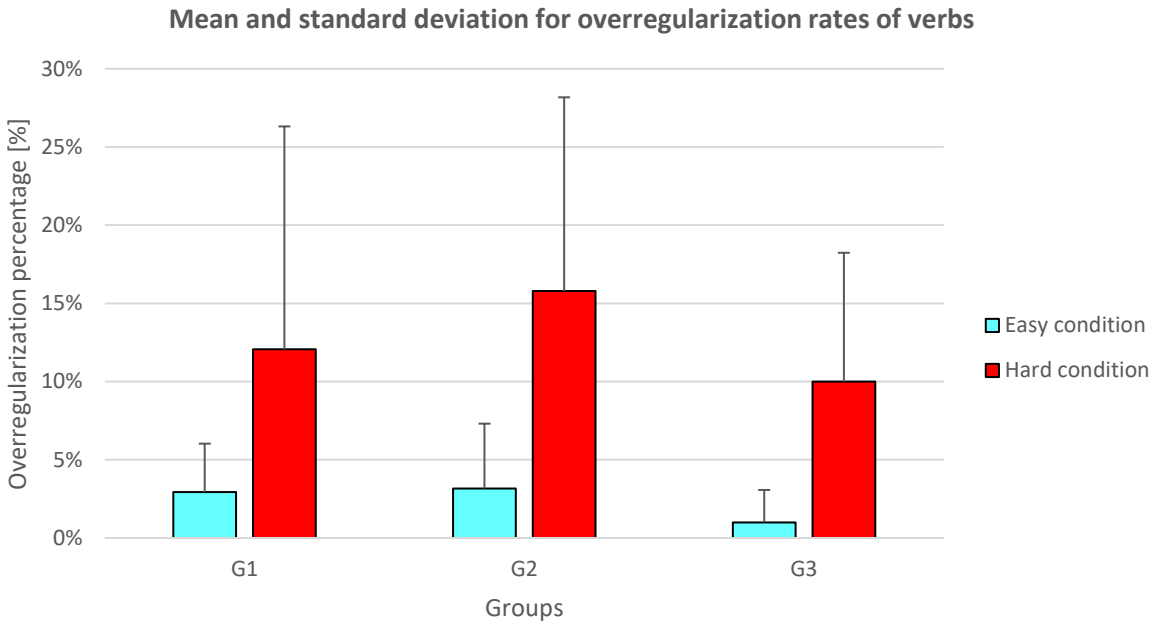


Figure 15: Comparison between groups for mean scores and standard deviation, overregularization of verbs, easy and hard condition.

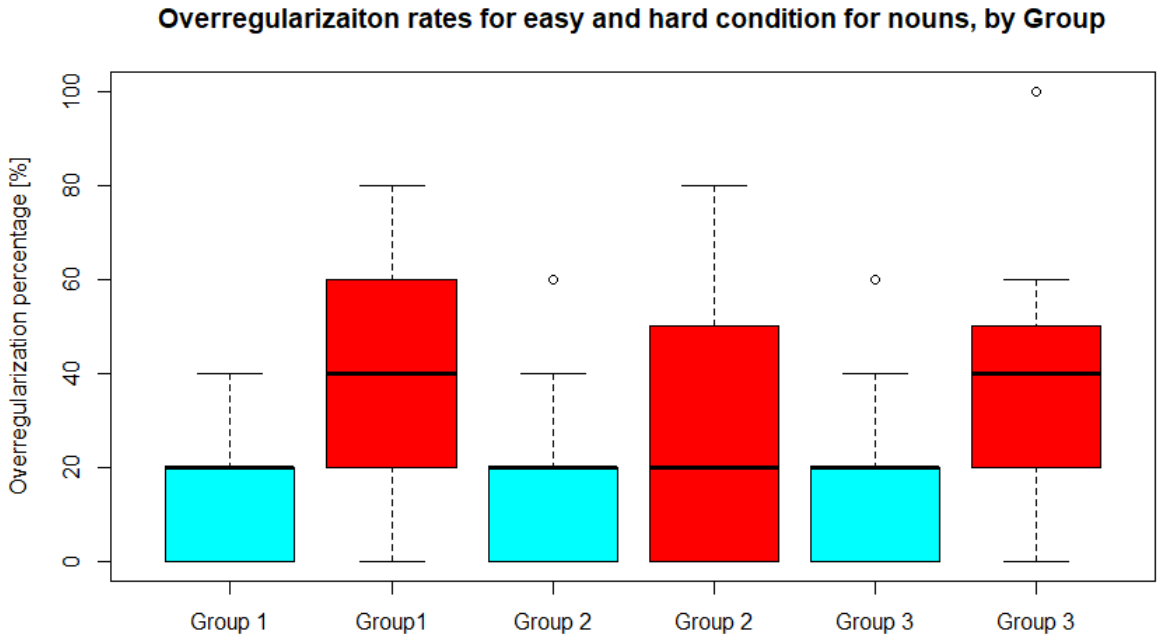


Figure 16: Comparison between easy condition (turquoise) and hard condition (red) for rate of overregularization of nouns, by Group.

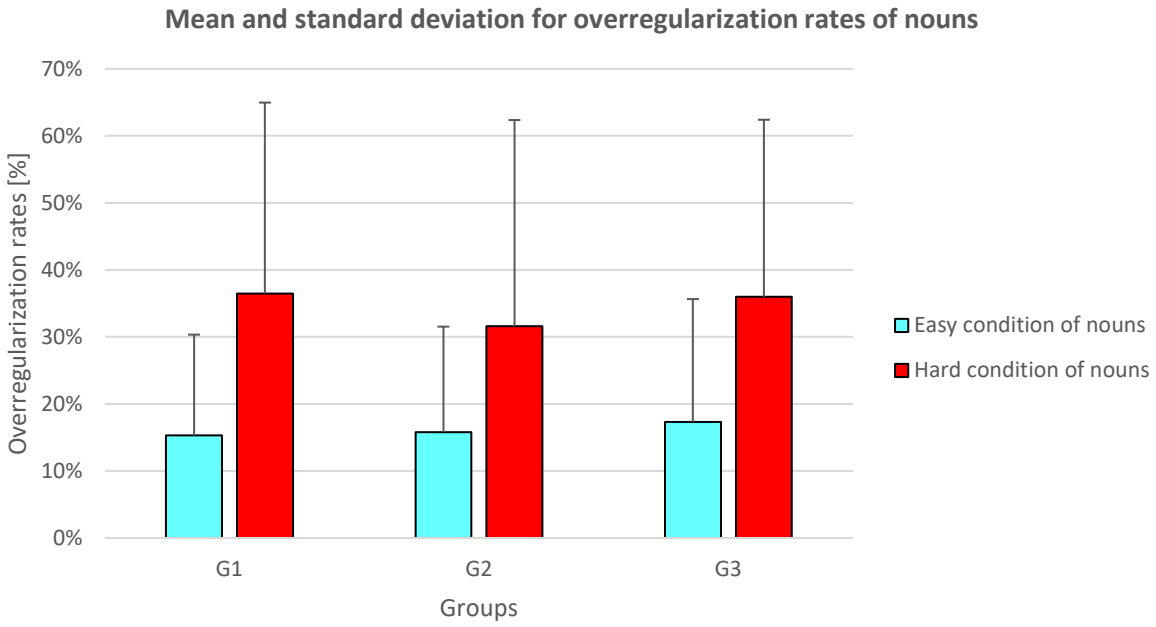


Figure 17: Comparison between groups for overregularization rate of nouns, easy and hard condition.

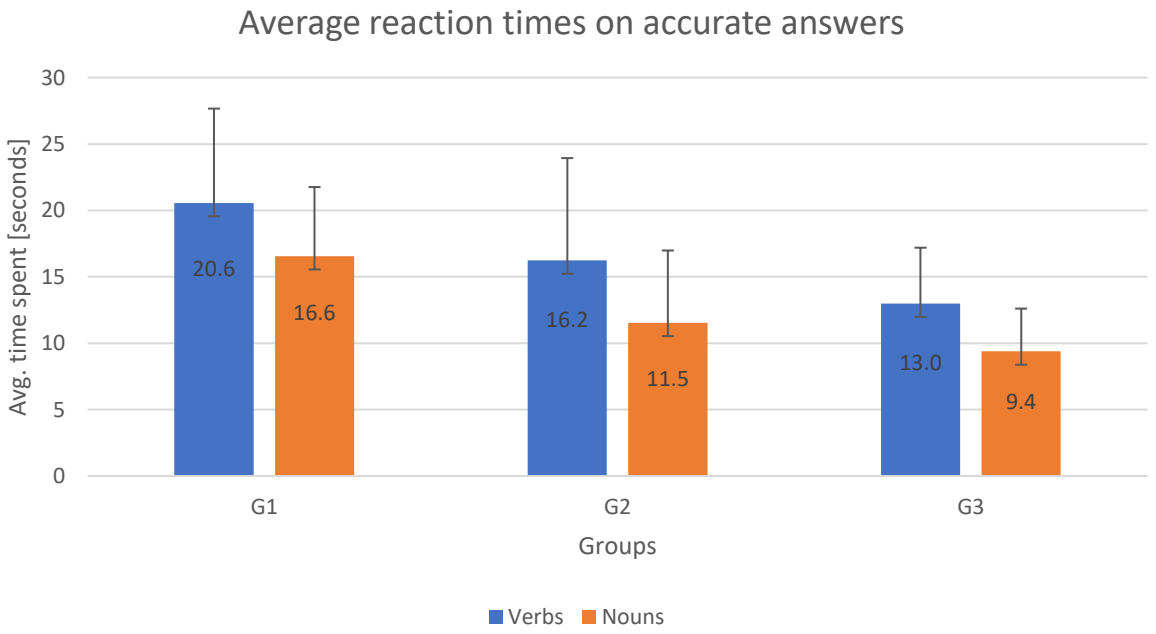


Figure 18: Average time spent on accurate answers with standard deviation error bars. Average time spent on accurate verb responses (blue) and for nouns (orange) are shown.

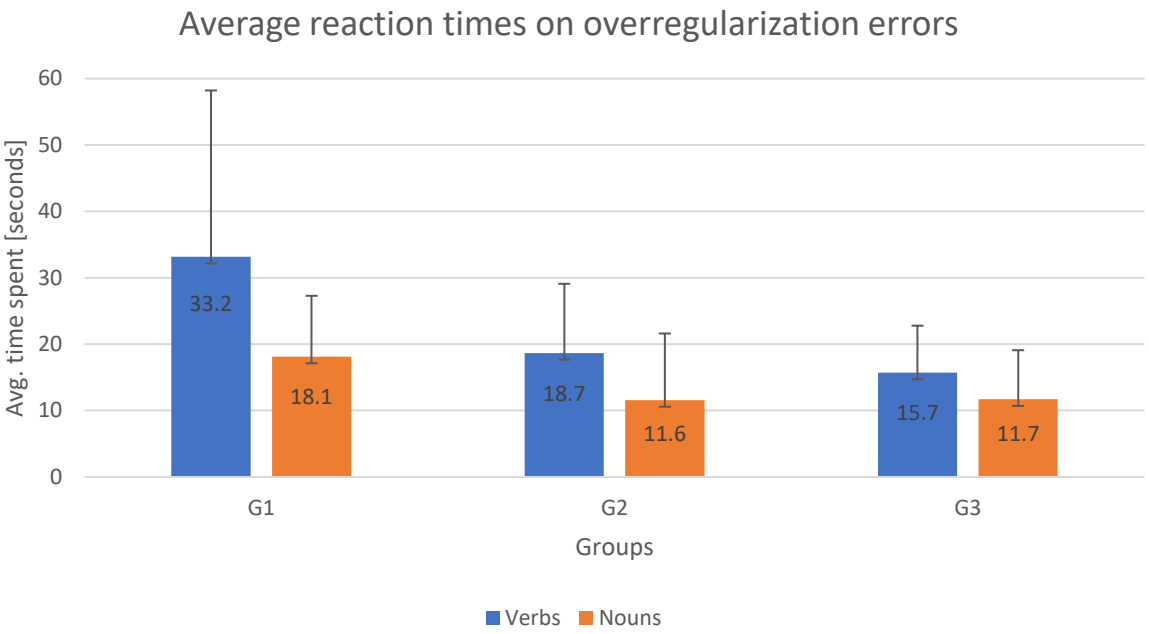


Figure 19: Average time spent on overregularization errors, including error bars for standard deviation. Average time on overregularization errors for verbs (blue) and for nouns (orange).

APPENDIX F

Tables

Shapiro-Wilk test for normality, verb accuracy			
	Group:	**W:**	**P-value:**
Overall rate of accuracy	G1	0.90045	p-value = 0.09669
	G2	0.9411	p-value = 0.276
	G3	0.83475	p-value = 0.01064
Easy condition for rate of accuracy	G1	0.96375	p-value = 0.7027
	G2	0.73515	p-value = 0.0001505
	G3	0.79585	p-value = 0.003254
Hard condition for rate of accuracy	G1	0.94704	p-value = 0.4114
	G2	0.94991	p-value = 0.3938
	G3	0.90045	p-value = 0.09669

Table 1: Shapiro-Wilk test for normality, accuracy distributions for verbs.

Wilcoxon test between groups for overall accuracy		
Groups being compared	**W:**	**P-value:**
Group 1 and Group 2	W = 137	p-value = 0.4454
Group 2 and Group 3	W = 66	p-value = 0.008083
Group 1 and Group 3	W = 48	p-value = 0.002734
Wilcoxon test between groups for accuracy, easy condition		
Groups being compared	**W:**	**P-value:**
Group 1 and Group 2	W = 123.5	p-value = 0.2246
Group 2 and Group 3	W = 98	p-value = 0.1133
Group 1 and Group 3	W = 62	p-value = 0.01185
Wilcoxon test between groups for accuracy, hard condition		
Groups being compared	**W:**	**P-value:**
Group 1 and Group 2	W = 144.5	p-value = 0.5979
Group 2 and Group 3	W = 65	p-value = 0.007152
Group 1 and Group 3	W = 53	p-value = 0.004987

Table 2: Results from group comparisons on accuracy rates. Result of an unpaired Wilcoxon-test.

Shapiro-Wilk test for normality			
	Group:	**W:**	**P-value:**
Overall rate of overregularization	G1	0.78067	0.001116
	G2	0.85953	0.009626
	G3	0.81933	0.00658
Easy condition for rate of overregularization	G1	0.75245	0.000485
	G2	0.71361	8.11E-05
	G3	0.49944	3.48E-06
Hard condition for rate of overregularization	G1	0.78956	0.001464
	G2	0.90579	0.06199
	G3	0.81184	0.005237

Table 3: Shapiro-Wilk test for normality, overregularization distributions.

Wilcoxon test between groups for overall overregularization rate		
Groups being compared:	**W:**	**P-value:**
Group 1 and Group 2	W = 127	p-value = 0.2719
Group 2 and Group 3	W = 195	p-value = 0.06649
Group 1 and Group 3	W = 141	p-value = 0.6142
Wilcoxon test between groups for overregularization rate in the easy condition		
Groups being compared:	**W:**	**P-value:**
Group 1 and Group 2	W = 165	p-value = 0.9161
Group 2 and Group 3	W = 180	p-value = 0.1198
Group 1 and Group 3	W = 171	p-value = 0.05497
Wilcoxon test between groups for overregularization rate in the hard condition		
Groups being compared:	**W:**	**P-value:**
Group 1 and Group 2	W = 120	p-value = 0.1875
Group 2 and Group 3	W = 181.5	p-value = 0.1721
Group 1 and Group 3	W = 124	p-value = 0.9073

Table 4: Wilcoxon test between groups for overregularization rates.

Self-rated level of proficiency against:	Group:	Correlation:
Overall rate of overregularization	Group 1	0.00171426
	Group 2	-0.2487454
	Group 3	-0.4861471
Easy condition for rate of overregularization	Group 1	-0.2362087
	Group 2	-0.3860624
	Group 3	-0.559017
Hard condition for rate of overregularization	Group 1	0.05121768
	Group 2	-0.1553049
	Group 3	-0.4214636

Table 5: Pearson correlation test between overall rate of overregularization for verbs and the participant's self-rated level of proficiency.

	Sum:	Mean:	Median:
8th grade	23	1.35	1
9th grade	32	1.68	1
10th grade	11	0.73	1

Table 6: Bare forms for verbs.

	Sum:	Mean:	Median:
8th grade	3	0.17	0
9th grade	1	0.05	0
10th grade	0	0	0

Table 7: Cases of approximation + overregularization for verbs.

	Sum:	Mean:	Median:
8th grade	8	0.47	0
9th grade	5	0.25	0
10th grade	5	0.33	0

Table 8: Bare forms for nouns.

	Sum:	Mean:	Median:
8th grade	2	0.1	0
9th grade	3	0.15	0
10th grade	1	0.06	0

Table 9: Cases of approximation + overregularization for nouns.

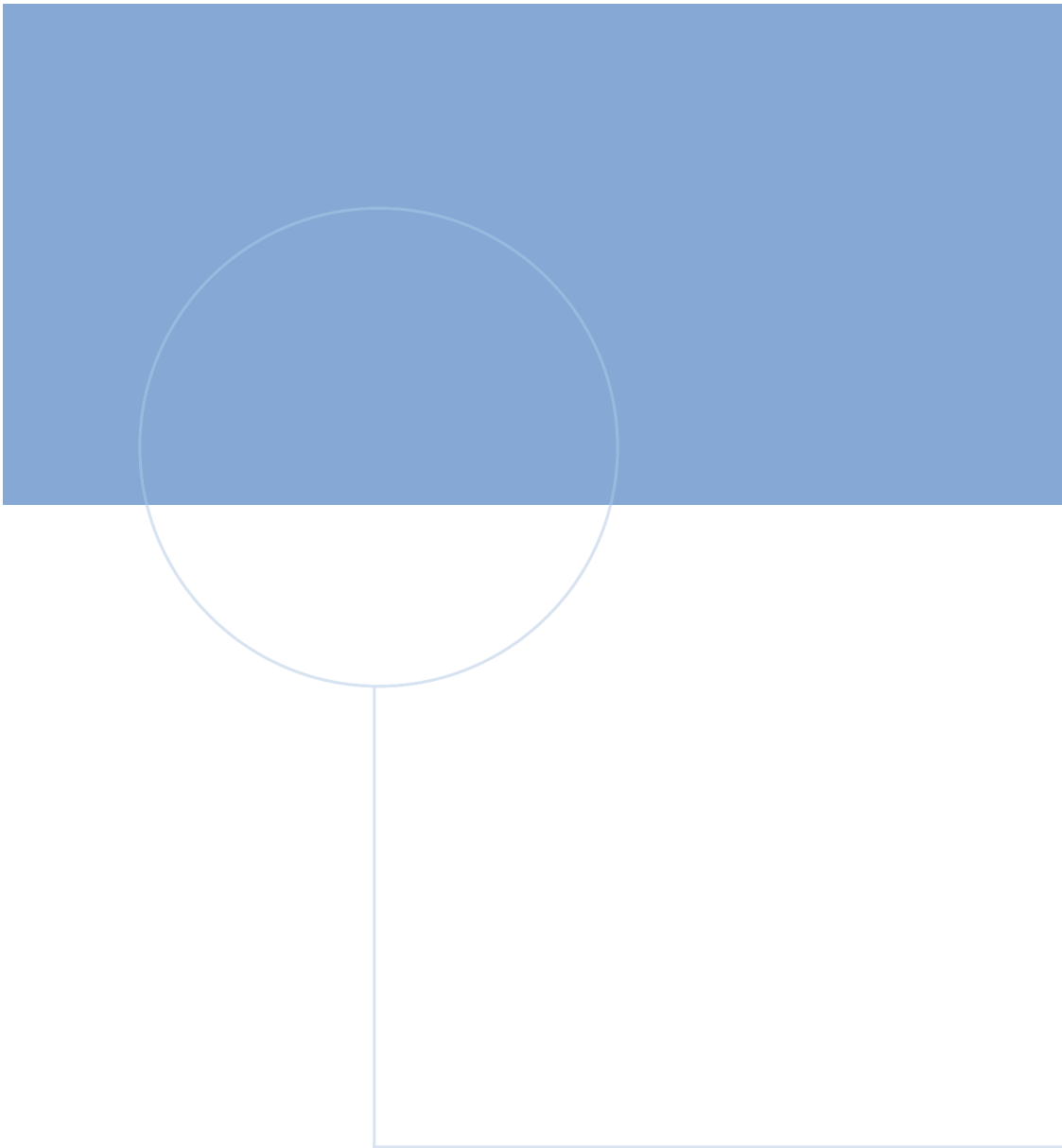
APPENDIX G

Relevance for teaching

This study examined whether we could find any evidence for U-shaped learning during L2 acquisition among English speaking Norwegian students in the 8th, 9th, and 10th grade. This phenomenon is of special interest for the organization of the mental faculty and the mental lexicon. Additionally, the study examined evidence for a dual-mechanism organization in the L2.

Examining and understanding how people acquire and develop their L2 is important for the teaching profession, at least regarding language teaching. A central part of language teaching involves guiding students towards fruitful development and providing them with constructive feedback by highlighting specific aspects of language that they should pay special attention to. Irregularities in tense inflection is one such aspect for instance. During conversations with some of the teachers involved in the data collection for this thesis, I learned that there are English teachers at the local schools who have noticed overregularization errors among students in lower-secondary schools. Anecdotally, one of them reported how these errors were persistent, especially the 8th grade, but that they later reduce in frequency. Conducting studies that explore this phenomenon in detail provides a clearer picture of how the students develop over time. I believe this study is of relevance for the teaching profession since it provides some descriptive stats related to accuracy rates, overregularization rates, as well as some descriptions of the other types of errors made by the students in the focus groups.

As mentioned in the discussion of this thesis, this study cannot confirm nor disconfirm the existence of U-shaped learning in the L2. However, if it is indeed a universal pattern that occurs in the formation of a new lexicon (as when acquiring a new language), then it might have implications for how teachers should approach and perhaps even view overregularization errors for instance. Given that this assumption is true, then this error type is indicative of a restructuring of the mental lexicon, which in turn is useful to know as a teacher. Further, if this kind of error is simply a natural consequence of the restructuring of the mental lexicon, then perhaps it is not necessary to heavily emphasize and correct it. That is not the same as stating that teachers should refrain from providing feedback on overregularization errors. Exposing the students to the normatively correct past tense forms is beneficial seeing as the frequency of exposure seems to have an impact on the correct formation of the irregular past tense. It might additionally help develop their metalinguistic knowledge.



NTNU
Norwegian University of
Science and Technology