Mali Kokvoll

L2 acquisition of agreement morphology: English subject-verb agreement among native Norwegians

Master's thesis

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

NTNU
Norwegian University of
Science and Technology

Mali Kokvoll

L2 acquisition of agreement morphology: English subject-verb agreement among native Norwegians

Master's thesis in Language Studies with Teacher Education
Supervisor: Dave Whitney Kush
May 2021

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

NTNU
Norwegian University of
Science and Technology

Abstract

Norwegian learners of English struggle with *subject-verb agreement*, often producing errors well into advanced stages. An *Acceptability Judgement Task* (AJT) was conducted in order to examine Norwegian speakers (N=28) intuitions about English subject-verb agreement constructions and how different factors affect their judgments. The design manipulated three factors: (1) Grammaticality, (2) Subject Number, and (3) Verb type, which led to a 2x2x2 factorial design with eight conditions. The results from the AJT were compared against the predictions of three different hypotheses that propose different explanations as to why L2 learners struggle with acquisition of functional morphology. The hypotheses were *Representational Deficit Hypothesis, Missing Surface Inflection Hypothesis and Feature Reassembly Hypothesis*.

The results revealed three main effects: (1) participants' average accuracy was marginally higher in grammatical conditions than in ungrammatical conditions, (2) error rates were higher overall when the subject NP was plural and (3) items with auxiliary verbs were rated more accurately on average than items with main verbs. In addition to this, two interaction effects were found showing that participants were less accurate with plural agreement when the verb was a main verb than auxiliary, and that errors with auxiliary verbs were rejected more consistently than errors with main verbs. Finally, the results also showed an interaction between self-reported hours of exposure to English and accuracy: participants' accuracy increased in line with how much English Media exposure they had per week.

The findings confirm that English subject-verb agreement is, in general, problematic for (advanced) Norwegian learners. The overall pattern of errors is not predicted by any of the hypotheses considered in isolation.

Acknowledgements

Table of Contents

Chapter 1: Introduction

This study is concerned with an area of English especially known to be problematic for second language (L2) learners, namely subject-verb agreement (SVA). English SVA is particularly interesting because it seems as L2 learners produce SVA errors even well into advanced stages (Garshol, 2019), and these errors seem to be independent of a learner's L1 as research shows that languages both with and without SVA have problems with this type of morphological dependency (Breiteneder, 2005).

Ample research has found that L2 learners produce both verbal and nominal agreement with varying success, where lexical items are both omitted and overproduced (e.g. White, 2003, p. 178; Ionin & Wexler, 2002, p. 95). More specifically, research on the acquisition of English SVA has shown that L2 learners frequently omit verbal inflection, where omission of the third person singular agreement marker -*s* normally is regarded as the most frequent error (Ionin & Wexler, 2002, p. 98; Breiteneder, 2005).

A question that has been raised is thus where this impairment of the use of Tense and Agreement morphology lies. Some researchers (e.g. Lardiere, 2009; Prevost & White, 2000) argue that the impairment is explained by the fact that the grammar or "interlanguage" of L2 learners contains abstract features and categories, and that the learners have problems with mapping the abstract features to the correct surface morphology. Other researcher, however, argue that the impairment lies in the representation of L2 functional categories themselves (e.g. Eubank, Bischof, Huffstutler, Leek & West, 1997). Other theories argue that difficulty with agreement may be largely a performance issue, not a grammatical issue (e.g. Slabakova, 2016; Prevost & White, 2000). This thesis explores these explanations of why L2 speakers seem to struggle with English SVA.

1.1 The present study

Based on this question, this study set out to investigate *intuitions of acceptability* with respect to the use of subject-verb agreement by Norwegian speakers of L2 English. This is further operationalized by looking at SVA in different configurations, and the aim of this paper is to explore

(i)   whether Norwegian L2 speakers of English accept or reject grammatical and ungrammatical subject-verb agreement constructions in general,

(ii)   what factors contribute to subject-verb agreement errors,

(iii)   whether the judgements of Norwegian speakers of L2 English are indicative of L1 influence, and

(iv)   if Norwegians' intuitions will align with predictions from different hypotheses regarding the source of difficulty with subject-verb agreement in L2.

To test Norwegian participants' intuitions about English SVA, an Acceptability Judgement Test (AJT) was conducted. The AJT manipulated three factors (Grammaticality, Verb Type, Subject Number), ultimately leading to a 2x2x2 factorial design in order to test predictions of three hypotheses about why L2 speakers make SVA errors: the Representational Deficit Hypothesis (e.g. Slabakova, 2016), Missing Surface Inflection Hypothesis (e.g. Prevost & White, 2000) and Feature Reassembly Hypothesis (e.g. Lardiere, 2009; Dominguez, Arche, Myles, 2011).

1.2 Overview

Before exploring the details of the AJT, an overview of relevant theoretical background is given in chapter 2. Chapter 2 first presents a comparative outline of agreement in both Norwegian and English. The chapter also provides general background on research in second language acquisition and what difficulties L2 speakers face in the acquisition of agreement morphology. Further, chapter 2 outlines the three hypotheses presented above, which offer different explanations for problems with functional morphology in L2.

Chapter 3 gives a brief theoretical background on the methodology used in the study. It also outlines the design of the AJT and a discussion of the predictions of the hypotheses for how participants will behave. Chapter 4 presents the results from the AJT. Further, these results are discussed in chapter 5 in light of both how the results align with the hypotheses and how they are in line with relevant theory on second language acquisition in general. Chapter 6 concludes the thesis.

Chapter 2: Theoretical background

This chapter begins with a comparison of subject-verb agreement in both English and Norwegian. Second, the chapter addresses theoretical issues in Generative second language acquisition (SLA) research and specifically acquisition of functional morphology. I offer an overview of difficulty L2 speakers have with agreement morphology and the errors they make. Finally, I talk about different theoretical accounts for why agreement can be challenging, which will be related to the specific case of Norwegian acquisition of L2 English agreement morphology, and the issues that arise in this specific case study.

2.1 Subject-verb agreement in English and Norwegian

Agreement is a wide-spread phenomenon which can be found in over 70 % of the world's languages, including English (Acuña-Fariña, 2012, p. 259). Although different theories explain subject-verb agreement in different manners (e.g. see Hudson, 1999), this study will focus on the "standard" view of English subject-verb agreement which holds that tensed verbs agree with their subjects in both person and number.

Subject-verb agreement is realized differently across verb types and tenses in English. Main verbs show fewer signs of overt agreement: most verbs in present tense do not have an agreement ending except for 3rd-person singular verbs which carry the -*s* suffix (see Table 1 for an overview). In the past tense, the form of a regular verb does not change depending on its subject. The auxiliary verb *be*, on the other hand, has even more distinct agreeing forms than main verbs in English (see Table 2). In addition to showing distinction in present 3rd person singular, *be* has different suppletive forms agreeing in person and number with the subject in present tense and number in past tense (Greenbaum & Nelson, 2009, p. 125).

Number: Person	Singular, *be*	Singular, *main verb*	Plural, *be*	Plural, *main verb*
1	am	love	are	love
2	are	love	are	love
3	is	loves	are	love
Number: Person	Singular, *be*	Singular, *main verb*	Plural, *be*	Plural, *main verb*
1	was	loved	were	loved
2	were	loved	were	loved
3	was	loved	were	loved

Table 1: Illustration of present and past tense *be*, and main verb *love*.

What separates English from Norwegian in terms of agreement is that Norwegian does not have any overt subject-verb agreement morphology (Holmberg & Platzack, 1995, p. 3). As seen in Table 2, main verbs like *elske* ('love') do not vary in form by person or number in either present or past tense. Finite forms of *be* are suppletive, but the conjugation is invariant across the paradigm (Garshol, 2019, p. 10).

Number: Person	Present, singular, *være*	Present, singular, *main verb*	Present, plural, *være*	Present, plural, *main verb*
1	er	elsker	er	elsker
2	er	elsker	er	elsker
3	er	elsker	er	elsker
Number: Person	Past, singular, *være*	Past, singular, *main verb*	Past, plural, *være*	Past, plural, *main verb*
1	var	elsket	var	elsket
2	var	elsket	var	elsket
3	var	elsket	var	elsket

Table 2*: Illustration of present and past tense *være* ('be'), and main verb *elske* ('love').*

2.2.3 The syntax of subject-verb agreement and affix lowering

To the extent that it is relevant, this thesis investigates the phenomenon of subject-verb agreement from a generative perspective. In this section I briefly introduce the basic syntax of the clause and subject-verb agreement. I adopt a standard analysis that divides a clause into three main domains: the Complementizer, Tense and Verb domains (Ramchand & Svenonius, 2014, p. 153; Slabakova, 2016, p. 212), which are organized in a hierarchy of C > T > V in all languages. In this structure, subjects are assumed to sit in the specifier of TP, where they can agree with the T head. Agreement morphemes occupy the T head, i.e. the agreement morpheme and the verb are distinct syntactic units at Deep Structure (Carnie, 2013, p. 220, 300). Agreement morphemes are their own heads, which surface in the T head, and they carry their own features separate from the verbs themselves. (See Figure 1).
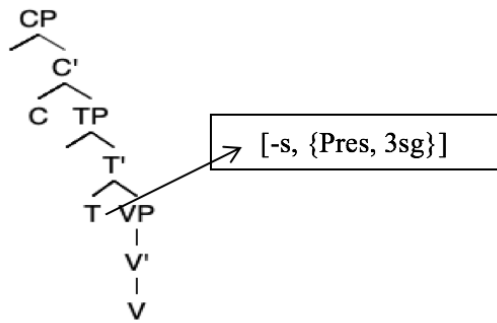
Figure 1: representation of clause domains and example of T head feature.

In this Deep Structure level which portrays the underlying representation of a clause, both auxiliaries and agreement morphemes appear to the left of the verb. However, agreement morphemes must be pronounced on verbs, so the lexical items in V and T have to come 'together' at some point. There are two ways this can be done. A verb can either raise to T, or the T morpheme can 'lower' onto V. A widely accepted fact in English is that main verbs do not raise to T. Instead of raising to T, lexical verbs somehow have finite inflections lowered onto them, known as *affix lowering* or *affix hopping* (Harwood, 2014, p. 295; Radford, 2004, p. 118; Carnie, 2013, p. 220) (See Figure 2).
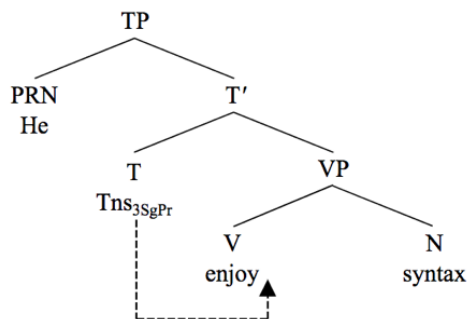
Figure 2: Affix lowering with lexical verbs (Radford, 2014, p. 118)

In contrast, finite auxiliary verbs *do* raise to T for agreement (Harwood, 2014, p. 295-296). Zobl & Liceras (1994, p. 163-165) write that the auxiliary *be* heads its own VP projection and raises to the T head. Suppletive forms of copula and auxiliary *be* are inserted to replace be+{T head with features}, created by raising and adjunction.

Tense morphology is assumed to be hosted in T in Norwegian, as in English. However, a difference from English is that Norwegian is a V2-language, i.e., the finite verb is the second constituent in a declarative main clause (Adger, 2003, p. 329), meaning that it has moved to C. It is assumed that the verb is moved to T before moving further to C. This means

that affix lowering is not needed in Norwegian main clauses since the verb picks up tense in T before moving to C. The syntactic and morphological differences between main and auxiliary verbs might be important to keep in mind when considering how Norwegian speakers of L2 English represent agreement in English.

2.2 Second Language Acquisition

Second language acquisition (SLA) is concerned with fundamental questions about learners' internalization of a second language and how to use this system in both speech production and comprehension (VanPatten & Benati, p. 2). This is a complex process which requires both effort, motivation and perseverance from the learner. However, there is no single approach which is universally accepted in the study of SLA, much due to its complexity and the fact that SLA falls within the scope of several already-established disciplines (Ritchie & Bhatia, 2009, p. 45).

An important aspect of SLA is the range of outcomes a L2 learner can have. White (2003, p. 241) writes that L1 acquisition is a process where typical learners essentially achieve the same end state with a steady rate of acquisition. In L2, however, we know that the end state of learners differs from native-speakers' and that there is significant variability in ultimate attainment among learners. Moreover, little is known about how to characterize the steady state L2 learners achieve (White, 2003, p. 241).

One widely accepted fact is that native language grammar seems to influence the outcome of the L2 language. White (2003, p. 45) writes that L2 learners never start with a clean slate, since their L1 grammar already is internalized in their cognition. It is thus reasonable to assume that some of this already-internalized grammar is being *transferred to*, or *interferes* with the L2 acquisition (Shimanskaya, 2015, p. 1; Ortega, 2013, p. 31). A central question both in SLA in general, but also in this study, is thus what is transferred from the L1 to the L2 grammar.

It is not my attempt to provide a comprehensive overview of theories within SLA research, however some main issues and how they relate to morphological acquisition in L2 are now presented.

6

2.2.1 Second language acquisition of morphology

The acquisition of functional morphology in L2 has been of great interest since the 1970s (Slabakova, 2016, p, 175-176). After attaining lexical items in an L2, the functional morphology becomes an important form-function mapping the learner has to acquire. Even if a learner has acquired several lexical items and can form comprehensible messages or sentences, their sentences may not necessarily be well-formed and grammatically correct (Slabakova, 2016, p. 175). A common observation is that L2 learners inconsistently use the correct morphology (morphology which carries information about gender, case, agreement, tense, number, etc.,) in their target language. Learners produce forms where appropriate morphology is present (1a, b), but also forms that use the incorrect morphemes (1c), or omit necessary morphemes entirely (1d) (White, 2003, p. 178).

(1)

        a. The girl plays piano

        b. The girl is playing the piano

        c. *The girl are playing

        d. *The girl play the piano

Since functional morphology typically will be represented differently in one's native language and target language, it is an area where L2 learners often struggle with acquisition.

Much research suggests that although L2 learners *are* inconsistent in their use of functional morphology, not all morphemes are equally hard to acquire for L2 learners (see 2). Bailey, Madden and Krashen (1974), looked at how accurate L2 learners of English with a wide variety of L1s were in their use of and accuracy with different functional morphemes (plural *-s*, progressive *-ing*, third person singular *-s* etc.). Their results indicated that the L2 learners acquired inflectional morphemes in ordered sequences. Results particularly interesting to this study was that learners made fewer mistakes and acquired the *-s* affix of plural nouns earlier than the 3rd person singular verb agreement morpheme *-s* (Bailey, et.al., 1974, p. 241; Slabakova, 2016, p.). This indicates that the learners do not necessarily struggle with acquisition of L2 functional morphology in general but rather with inflectional morphology for English SVA.

(2)

  a. *-s* affix, plural nouns: *"The boy threw the balls"*

  b. *-s* affix, 3rd person singular: *"The boy throws the balls"*

Research on acquisition of L2 inflectional morphology shows that it is more common for L2 speakers of English to neglect the use of inflection rather than to use the wrong verbal inflection (Garshol, 2019, p. ii). However, research on Norwegian speakers of L2 English shows that Norwegian learners commonly overproduce the 3rd person singular marker in contexts where it should not be present (Garshol, 2019, p. 74). Norwegians also make omission errors, but Garshol (2019) writes that omission errors and overproduction errors differ based on the type of subjects. The majority of omission errors happen in clauses with personal pronouns as subjects (3b), while overproduction errors happen in clauses with NP subjects (3a) (Garshol, 2019, p. 74).

(3)

  a. Overproduction error: *The dogs runs fast, [NP + T {pres, 3, pl} + V + Adv]

  b. Omission error: *He run fast, [PP + T {pres, 3, sg} + V + Adv]

Garshol (2019) suggests that Norwegian speakers of L2 English consider the marked affixal form of the finite verb a default finite form which they resort to when the context is complex, for example complex NPs as subjects. Personal pronouns are seen as somewhat simpler, both syntactically and semantically, and Norwegian L2 learners of English thus make fewer errors in total with these clauses (Garshol, 2019, p. 75). The next section will address another inflection paradigm in addition to affixal inflection – suppletive inflection.

2.2.2 Suppletive versus affixal inflection in second language acquisition
English has tense morphemes consisting of both bound and unbound morphemes. The bound morphemes are suffixes such as third person singular *-s* and past tense *-ed*, which affix to regular verbs. Suppletive agreement occurs with irregular verbs like the *be* copula, auxiliary *have* and other irregular verbs (Mayo; Olaizola, 2010, p. 132; Ionin & Wexler, 2002, p. 102). This study only considers copula and auxiliary *be* and third person singular *-s*.

Research on inflection in L2 has almost exclusively been based on affixal morphology, with focus on suffixes such as past tense *-ed* and the 3rd person singular *-s*, while suppletive agreement has often been disregarded (Ionin & Wexler, 2002, p. 102). An example

8

is Eubank, et. al.'s (1997) research, which looked at L2 learners' knowledge of verb raising and inflection. The study only looked at the participants ability to produce correct agreement with affixal agreement, more specifically their ability to correctly produce *-s* in translation tasks (Eubank, et. al., 1997). Items with the copula *be* were only used as a distractor in this study, and not taken into consideration. Disregarding suppletive agreement might lead researchers to underestimate the knowledge L2 learners have of inflection in general or might obscure general problems with general agreement that go beyond regular agreement morphology (Ionin & Wexler, 2002, p. 103; Lardiere, 1999).[1] Since little research has been done on suppletive agreement it will be an interesting aspect to further investigate.

2.3 The Role of L1 in SLA

I now proceed to cover some theoretical background within SLA that is relevant for my study. Section 2.3.1 addresses the widely accepted idea that L1 influences the acquisition and perhaps stable state of L2. Later sections examine specific proposals about the acquisition and representation of functional morphology in L2. These theories all offer views on the issue of L1 influence and the question of what may or may not cause difficulties for the Norwegian L2 learner when looking at subject-verb agreement in English.

2.3.1 First language transfer

A central question in SLA is how much and what parts of their L1 grammar L2 learners bring to the L2 acquisition process. White (2003, p. 45) writes that there exists almost universal consensus that L2 learners do not start with a clean slate since their native grammar already is internalized in their cognition. It is thus plausible that the L1 will influence the L2 in some way or another, but the central question is exactly where and how the already internalized L1 grammar is *transferred to, influences* or *interferes with* L2 grammar and processing (Shimanskaya, 2015, p. 1; Ortega, 2013, p. 31).

As a point of terminology: generative SLA research assumes that L2 learners construct and use a rule-based system for their L2. The L2 grammar may have properties of both L1

---

[1] Some research has been conducted on suppletive agreement. This research presents evidence of both copula *be* and auxiliary *be* being acquired before suffixal agreement endings (Ionin & Wexler, 2002; Zobl & Liceras, 1994). Zobl & Liceras (1994, p. 174) further argue this to be an indication of functional projections being available early in L2 acquisition and instantiated through the use of different *be* forms.
Garshol also looked at suppletive agreement among Norwegians and found that Norwegian learners of English often produce incorrect suppletive agreement, where they tend to overproduce the plural forms of the verb *be* (2019, p. 180)

and L2, so it is sometimes referred to as *interlanguage* (Shimanskaya, 2015, p. 3). Interlanguages are often thought to be coherent, rule-based systems.[2]

Previous studies have shown that interlanguages are influenced both by native language (L1 transfer) and properties of the second language itself. Though there is varying agreement on what transfers and not, there seems to be consensus that it is "doubtful that there are grounds for dismissing, or at least disregarding, the notion of L1 influence entirely" (Wold, 2017, p. 38). Accepting this claim, however, it still needs to be established *to what extent* the L1 will influence a learner's developing L2 grammar. There are several theories which address this issue, one of them being Schwartz and Sprouse's (1996) *Full Access/Full Transfer Hypothesis.* Their hypothesis claims that the "initial state of L2 acquisition is the final state of L1 acquisition" (Schwartz and Sprouse, 1996, p. 40-41), i.e., the L1 grammar fully transfers in the early stages of the L2 acquisition process. However, as the learner gains positive evidence from L2 input, the initial state will change and the learner will be able to restructure their L2 grammar (Schwartz and Sprouse, 1996).

2.3.1.1 Are L2 grammars "complete" grammars?

Some theories hold that the L1 plays a permanently constraining role on L2 acquisition and that the interlanguage does not have all of the characteristics of a 'full' grammar. Some accounts question whether L2 grammars/systems/interlanguages are "grammars" in the same way as L1 systems. These issues are embodied within the Representational Deficit Hypothesis (RDH) (e.g. Schwartz and Sprouse, 1996). RDH holds that L2 grammars may be 'missing' some features of full grammars, that are observed in L1. The hypothesis states that abstract morphosyntactic features are present in L2 interlanguages *only* if the same abstract morphosyntactic features also are represented in their L1 (Slabakova, 2016, p. 186). It thus becomes clear that the two issues regarding transfer and whether or not L2 grammars are "grammars" in the same way as in L1 interact.

2.3.2 Theories of Morphological Difficulty

As shown in previous sections, functional morphology is acquired at a variable rate and with varying success. Different theories have thus been presented in order to examine problems

---

[2] Given the idea of interlanguage being a coherent system, an important point to note is that research may not always characterize observed deviations from the target language as "errors" (Wold, 2017, p, 34). Instead they reflect the output of a different system from the target language, and the scientific study of properties of *interlanguage* in its own right is a major area of focus in SLA.

regarding acquisition of functional morphology. This subsection presents some of these theories in order to see how they explain these issues.

2.3.2.1 Representational Deficit Hypothesis

The Representational Deficit Hypothesis (RDH) is concerned with the underlying knowledge of syntactic representation. The RDH holds that poor morpho-syntactic performance in L2 is due to incomplete grammatical competence. Researchers supporting this theory argue that abstract morphosyntactic features, for example those relevant for agreement, are accessible to adult L2 learners only if the same morphosyntactic features are represented in their native language. Thus, the L2 learner is dependent on having the same morphosyntactic features in their own native language in order to perform well in their L2 (Slabakova, 2016, p. 186, 187). This will ultimately mean that if a L2 learner's native language does not have a certain morphosyntactic feature that is used in their L2, use of this morphosyntactic feature will always be impaired in their L2. An example of this is seen when comparing Norwegian and English and which features are present in T. English is assumed to have person and number features on T, which are relevant for agreement, while Norwegian does not have these features on T. According to RDH, Norwegians are therefore not predicted to have person and number features on T in L2 English, as seen in (4), where (4a) illustrates L1 English and (4b) illustrates the Norwegian L2 prediction.

(4)

      a. Mary play**s** [T {pres, 3, sg} + V]   (-*s*)
      b. *Mary play [T {pres, ~~3, sg~~} + V]   (Ø)

There are, however, ways in which L2 learners might display successful performance on morphosyntactic tasks like agreement, even if the specific morphosyntactic feature is not present in the learners' native language. Supporters of the RDH argue that successful performance might happen when L2 learners notice different morphemes and imitate native speakers, or "fake" correct performance (Slabakova, 2016, p. 187). This notion of imitating native speakers causes L2 learners to create rules about when to use which form of the verb, and these rules will be strengthened with increased exposure to the target language.

2.3.2.2 Missing Surface Inflection Hypothesis

The Missing Surface Inflection Hypothesis (MSIH) proposes that variability in adult L2 performance with functional morphology reflects difficulties with overt realization of morphology rather than a deep problem with the underlying syntactic representations speakers need in order to produce correct inflectional endings (Prevost & White, 2000, p. 104, 108; Slabakova, 2016, p. 190). The hypothesis proposes that L2 learners have unconscious knowledge of the features underlying both tense and agreement, but struggle with realization of correct inflectional endings in production. It is therefore a mapping problem between correct abstract features and morphological form (Slabakova, 2016, p. 193). In other words, problems with subject-verb agreement occur because L2 learners struggle with retrieval of the relevant lexical items that are necessary for successful inflection, meaning that MSIH is a *performance hypothesis* rather than a hypothesis concerning *competence*.

MSIH proposes that the link between morphology and syntactic knowledge is arguably not strong enough to guarantee reliable production even though the learner has acquired the syntactic features of a functional category. An important question is thus what grammatical process allows this kind of separation between a realized morphological form and unimpaired syntactic knowledge (Slabakova, 2016, p. 191). Slabakova states that it is necessary to look at *Distributed Morphology*, and more specifically its claim of lexical insertion. In Distributed Morphology there is no divide between the construction of words and sentences (Slabakova, 2016, p. 191; Prevost & White, 2000, p. 127). Each inflected form is associated with several features like person, number, tense and gender and for lexical insertion to happen, there must be consistency between the features of lexical items and the features of the syntactic node where it should be inserted (see 5).

(5)

      a. [T {pres, 3, sg}]   (-s)
      b. [T {pres, 1, sg}]   (Ø)

The assumption is that L2 learners have acquired the relevant features of the syntactic node, but the learners' impairment lies in the ability to retrieve the correct form to insert (Prevost & White, 2000, p. 127). Instead of retrieving the correct lexical item, learners can retrieve more 'accessible' forms, for example "default forms" such as bare forms/infinitives. These default forms are often more accessible because of how basic they are and their frequency.

In summary, the Missing Surface Inflection Hypothesis tries to offer an explanation of the relationship between functional morphology and knowledge of the syntactic properties that underlies this functional morphology. In a nutshell one can describe the main proposal in MSIH as a mapping problem between syntax and morphology – learners have the syntactic knowledge needed to retrieve the correct lexical item and perform correct inflection, but sometimes learners retrieve the wrong lexical item because other items are more accessible.

2.3.2.3 The Feature Reassembly Hypothesis

Recent research within the Generative Grammar framework has treated abstract linguistic features as "the basic unit of currency for describing differences between languages (Lardiere, 2009, p. 180). These abstract linguistic features carry phonological, semantic and formal information, and differences in these features is what causes differences between languages (Lardiere 2009; Slabakova, 2016, p. 197). Cross-linguistic variation is thus described in terms of what features are selected from the universal inventory (Universal Grammar, or UG) in addition to how these chosen features are assembled into lexical items specific for each language (Lardiere, 2009, p. 189; Dominguez et. al., 2011, p. 183) (See Figure 3).

Language Acquisition = (Feature Selection + Feature Assembly)

Feature Selection: $F \xrightarrow{CH_L} [F_{L1}]$

Feature Assembly: $[F_{L1}] \longrightarrow [Lex_{L1}]$

Figure 3: The process of language acquisition in L1   (Dominguez et. al., 2011, p. 183)

When acquiring a L1, children need to select only that subset of features that is required to explain patterns in the input of the specific language being acquired, at the same time as they "discard" others (Lardiere, 2009, p. 174).

The selection of these features is, however, different in L2 since L2 learners bring "an already-fully-assembled set of (L1) grammatical categories" which correspond to different bundles of features (Lardiere, 2009, p. 175). The FRH thus argues that "during the second reassembly stage, L2 learners are expected to implement several modifications to the featural organization that they transferred from L1" (Slabakova, 2016, p. 76). As such, errors in L2 are explained by challenges with reconfiguration of the way features are organized because features from L1 might use different lexical items than the same features in L2.

An example of the challenges of feature reassembly can be seen when comparing Norwegian and English, and what features that are present in T heads. In English, present tense subjects need to agree with their verbs – i.e. English T heads carry Tense (+pres, +past, etc.), person (1st, 2nd, 3rd) and number features (sg/pl). As seen in (6c, d), the different morphological items are associated with different feature bundles. The features vary with the feature of the subject. In Norwegian, however, present tense main verbs are marked with the affix *-er* in sentences with both singular and plural subjects, so *-er* does not have any person or number features (see 6).

(6)

      a. Hun skriv-**er**   (-er) [+pres]

      b. De skriv-**er**   (-er) [+pres]

      c. She write-**s**   (-s) [3rd, sg,+pres]

      d. They write-Ø   (-Ø) [3rd, pl, +pres]

FRH assumes Full Transfer: that the L1 is the starting point, and forms the basis for acquisition of the L2 grammar system (Lardiere, 2009, p. 191). When L2 learners are exposed to target-language input, they look for matching forms in the L1 to those in the L2, and then *transfer* the feature combinations that are relevant onto the "L2 equivalents" (Shimanskaya, 2015, p. 22). Thus, Norwegians will try to use the L1 analysis of syntactic heads and their feature combinations in L2. Norwegians are expected to take the T suffix, -s, in English and assume that the head has the same features in L2 as it has in L1 (see 6). Since *-er* is only marked for [+pres] in Norwegian, Norwegians are predicted to treat the English *-s* morpheme as a simple marker of present tense (Garshol, 2019).

Dominguez et. al. (2011, p. 184) write that successful SLA thus "depends on whether L1 features have the same morpholexical expression in the L2 and whether learners can effectively reconfigure them when they do not". Since, as seen in (6), features from L1 might be on different lexical items in the L2 than they were in the L1 (Slabakova, 2016, p. 198), transfer will lead to cases where reassembly or reorganization of existing features is necessary. Thus, a challenge that L2 learners face is the 'reconfiguration' of the way features are organized and represented in the L1 into new configurations of lexical items in the L2 (Lardiere, 2009, p. 175) – e.g. Norwegians have to understand how to put new features [person, number] on English T heads.

14

2.4 Concluding section

This chapter has presented different theories within SLA which offer explanations of the causes for problems with functional morphology, which will lay as a theoretical backdrop for this study. The following chapter will investigate which of the theories presented above that can provide good explanations for why the acquisition and use of agreement morphology are challenging for L2 Norwegian speakers of English.

Chapter 3: Method

This study examines how Norwegian speakers of L2 English process English subject-verb agreement and their knowledge of English agreement in general. The study does so by looking at their ability to identify correct agreement and to detect English subject-verb agreement errors. To explore this, a binary, forced-choice *Acceptability Judgement Test* (AJT) was employed. This chapter first presents some theoretical background for choosing AJT. Secondly, the experiments participants will be explained, before describing the experimental design. Finally, the three hypotheses and their predictions are presented.

3.1 Acceptability judgement

It is important to choose a suitable study design in order to get valid evidence to support or discard different hypotheses. This study set out to gather data about intuitions L1 Norwegian speakers have regarding English subject-verb agreement. An AJT was chosen to measure participant intuitions. This chapter presents some theoretical background for this choice.

In an AJT, participants are asked to rate the *acceptability* of a sentence. The underlying assumption is that "*acceptability* is a *percept* that arises (spontaneously) in response to linguistic stimuli […]" (Schütze & Sprouse, 2012, p. 3). Acceptability is a percept, similar to any other perception like temperature, pain etc.. Since there is no direct way of measuring percepts, experiments depend on indirect measurement methods (Schütze & Sprouse, 2012, p. 3). One of the most common ways of indirectly measuring perception is to ask participants to report their perceived acceptability on a given scale. Acceptability judgements thus report perceptions, which makes AJs a type of behavioral response which most likely requires a cognitive explanation (Schütze & Sprouse, 2012, p. 3). When asking participants to observe their language perception consciously, i.e. from a metalinguistic perspective, the acceptability ratings are the results of conscious attitudes towards a specific sentence type.

There are several possibilities as to how judgement data can be collected. The experiment presented in this paper uses a binary yes-no acceptability task which asks the participants to judge whether a sentence is correct or not. The primary advantage of using this type of non-numerical task is that it is quick to deploy. Secondly, yes-no judgements can be used to compare the relative difference between conditions, which is done by computing the proportion of yes-responses for each condition (Schütze and Sprouse, 2012, p. 6).

However, all research designs can potentially create a few challenges, which also is true for AJTs. The first thing to consider is the fact that the items often are presented in isolation and not in context, which might have yielded different responses. Another issue is the fact that there is no way of knowing the reason behind why a participant judges a sentence as "good" or "bad". The participants might thus reject an ungrammatical sentence due to other factors than SVA errors. However, by presenting multiple items from the same condition (as done in this experiment), this might help control some of the challenges presented above, because we expect an overall behavioral trend to emerge in the *average* response to a condition.

3.1.1. Grammatical versus ungrammatical sentences in AJTs

In the present study, participants in the AJT are asked to rate both grammatical and ungrammatical sentences. Judgement data is most commonly used to measure knowledge of form and to understand the type of knowledge L2 speakers hold (Spinner & Gass, 2019, p. 31). The question is then why we want participants to rate both acceptable and unacceptable items.

The first thing to consider is the question about optionality – determining whether participants are at chance in their performance. If the study set out only to test acceptable sentences, it would be problematic to determine whether participants actually *also* accepted the unacceptable versions. Optionality is thus one reason for including both acceptable and unacceptable items. Also, in contrast to for example corpus analysis, linguistic judgments might reveal what structures in a language are disallowed (Huang & Ferreira, 2020). Acceptability judgments thus allow researchers to test predictions in terms of what forms are generated or not in a grammar. Moreover, the implementation of both grammatical and ungrammatical sentences allows us to measure "bias" – i.e. how likely participants are to say *yes* or *no* to a sentence. Since acceptability is seen as a percept, there can be some bias in responses. Participants might for example be biased to say *yes* to an unacceptable sentence in cases they are unsure about, while other participants might be very strict and rate every sentence that sounds somewhat odd to them as unacceptable (Huang & Ferreira, 2020). It is therefore necessary to implement multiple items for the same condition and in addition to this, the number of grammatical and ungrammatical items should also be equal across the study. A greater number of for example grammatical sentences can lead the participants to expect all test items to be grammatical and influence their judgements in general (Huang & Ferreira, 2020; Shütze,1996).

18

3.1.2 Explicit versus implicit knowledge

The distinction between explicit and implicit knowledge is significant when discussing L1 and L2 acquisition. Children acquire their L1 by interacting with their caretakers in natural settings and with natural communication. Through this interaction they automatically acquire complex knowledge of the structure in their L1 (Ellis, 2008, p. 1). Paradoxically, children cannot describe their knowledge, i.e., their knowledge is implicit (Ellis, 2008, p. 1: Rebuschat, 2013, p. 593). As such, acquisition of L1 grammar is considered implicit since L1 learners extract their knowledge from natural settings rather than explicit rules. Adult L2 acquisition, on the other hand, is often a combination of implicit and explicit learning. There are often limitations as to what L2 learners might acquire implicitly from communicative contexts. Thus, additional resources of explicit learning are often necessary for adult attainment of L2 accuracy (Ellis, 2008, p, 1).

Participants in this study have most likely acquired their L2 through a combination of both explicit and implicit language learning. Ellis (2009, p. 15, 27) writes that there is a problem in determining whether explicit and implicit knowledge stores are separate or linked, since we cannot precisely determine how L2 learners draw on their linguistic knowledge when exposed to a language task. E.g. it might be the case that participants have developed implicit and explicit knowledge of the *same linguistic feature*. Learners might have internalized the verb [runs] (as a single item) as explicit knowledge, while the procedure for attaching the affix *-s* is internalized as implicit knowledge (Ellis, 2009, p. 15). It is thus difficult to precisely declare if the AJT in this study is testing implicit or explicit knowledge, which is also the case in SLA research in general. Spinner & Gass (2019, p. 32) write that "in current research […], the issue of the type of knowledge that is tapped by judgement data is quite controversial". However, Ellis (2009) *does* present several critical features which can distinguish measures of explicit vs. implicit knowledge. Among them is timed vs. untimed grammaticality judgement tests. The AJT in this study presented sentences in a timed phrase-by-phrase manner. However, there was no time limit as to how fast they needed to judge whether the sentence was acceptable or not. It is thus plausible to argue that AJTs test both explicit and implicit knowledge.

3.1.3 Performance versus competence

Another important distinction in second language research is *competence* versus *performance*. Chomsky introduced these terms in 1965, and he described competence as "the speaker-hearer´s knowledge of his language" and performance as "the actual use of language

in concrete situations" (1965, p. 4). Performance can thus be described as observable linguistic behavior, while competence refers to the description of abstract linguistic knowledge. All judgements tasks will always be filtered through performance in some way or another, i.e. acceptability judgments involve some sort of performance, and performance might include different confounding variables (Spinner & Gass, 2019, p. 16).

Chomsky further writes that "the notion of "acceptable" is not to be confused with "grammatical"" (1965, p. 11). Grammaticalness belongs to the study of competence and is an abstraction. As such, we cannot test competence directly, we can only draw inferences about it from speaker-hearer´s performance (Cowart, 1997, p. 6-7; Chomsky, 1965, p. 11; Tremblay, 2005, p. 133). Consequently, an acceptability judgement task do not give us direct access into linguistic competence, but is rather a tool which helps us make inferences about linguistic knowledge.

Acceptability and grammaticality should be seen as separate, but interrelated concepts. Acceptability percepts are influenced by performance factors separate from the grammar, which means that performance factors will have an impact on how participants rate sentences. For example, sentence (7) is grammatical, but it typically receives a low acceptability rating, because it is difficult to process. By contrast, sentences like (8) might score higher on acceptability, even though they are ungrammatical. (8) has an argument ("the student") that is not assigned any thematic role because the sentence is missing a verb. Nevertheless, research has shown that these sentences with the "missing verb effect" have been correlated to higher acceptability rates (Gibson & Thomas, 1999, p. 225).

(7) The student the teacher the school had hired thought met with the headmaster.
(8) *The student the teacher the school had hired met with the headmaster

Even though we cannot tap directly into grammatical competence with acceptability judgement tasks, it does not imply that acceptability judgements are not important to linguistic research. However, it means that when dealing with acceptability judgements, it is essential to search for systematic judgement patterns. Judgement data also allows us to look at the *processes* of grammar, which we have seen are believed to be inaccessible – and therefore only accessible through indirect study approaches (Schütze & Sprouse, 2012, p. 28). Grammaticality is thus thought to play an essential role in judgement of sentences, and acceptability judgements provide valuable information of linguistic intuitions that is not necessarily available from other kinds of data (Schütze & Sprouse, 2012, p. 29).

3.2 Participants

One participant group was recruited for the study: a group of L1 Norwegian speakers with English as their L2 (N= 28, 17 Female, 9 Male, 1 Other). Optimally, a control group consisting of native speakers should also have been recruited in order to see how L2 Norwegian speakers of English performed compared to the native speakers. However, since I experienced some problems in the process of recruiting native speakers, the study does not include a control group. The age span of the participants was between 21-42 years (Mean age= 26.4). The participants were exclusively recruited through public posts on social media with Facebook as the primary source.

Explicit consent was given by all the participants before completing the AJT (see Appendix II for the text of the consent form). Before the AJT, participants were asked to provide general information about their linguistic background and other factors which might have affected the results (see Appendix III). The processing of the participants' personal data was assessed to be in line with data protection legislation according to NSD.

3.3 Experimental design

3.3.1 Procedure

An Acceptability Judgement Test and a Confidence Rating (CR) were used. Data was collected using the online experiment platform called IbexFarm (Drummond, 2013). The study was made accessible by link.

After completing the consent- and background information form, the participants were presented with instructions as to how the experiment would work and how they would indicate acceptability and not. Participants were also asked to complete a couple of test sentences in order to confirm their understanding of the instructions.

Experimental sentences were presented one at a time. Each sentence was presented phrase-by-phrase using *Rapid Serial Visual Presentation* (RSVP) format. The original design of the experiment displayed a new phrase every 400ms, but after a pilot test this was increased to 450ms per phrase in order to give participants enough time to read the phrases fully. After each sentence, participants were to press "1" if they judged the sentence to be "good" and "0" if they judged it "bad". After making their acceptability judgment, participants rated their confidence in their judgment on a 3-point Likert-scale (*not confident*, *sort of confident*, *very confident*). A "Latin Square" method was used for randomization of the test items across participants.

3.3.2 Data materials: Test sentences

Eight conditions were constructed in order to test the hypotheses (see section 3.4 for discussion of hypotheses). The AJT items are accounted for in section 3.3.2.1 while a brief justification for choosing to include Confidence Rating is presented in section 3.3.2.2.

3.3.2.1 AJT test items

The AJT tested under what conditions Norwegian speakers of L2 English struggle with subject-verb agreement. The design manipulated three factors: (i) number of the subject [*singular* versus *plural*], (ii) grammatical agreement [*grammatical* (agreement with subject) versus *ungrammatical* (no agreement with subject)], and (iii) verb type [*auxiliary* versus *main verb*]. This lead to a 2x2x2 factorial design with eight conditions (see 9; outline of AJT design)

(9)

    a. The guest that arrived yesterday is complaining - [Sg Subj – Aux – G]

    b. The guest that arrived yesterday are complaining - [Sg Subj – Aux – UnG]

    c. The guests that arrived yesterday is complaining - [Pl Subj – Aux – UnG]

    d. The guests that arrived yesterday are complaining - [Pl Subj – Aux – G]

    e. The guest that arrived yesterday complains - [Sg Subj – Main – G]

    f. The guest that arrived yesterday complain - [Sg Subj – Main – UnG]

    g. The guests that arrived yesterday complains - [Pl Subj – Main – UnG]

    h. The guests that arrived yesterday complain - [Pl Subj – Main – G]

All test items contained a relative clause (RC) between the subject noun and verb. The RC verb did not agree with the subject to avoid giving cues to grammatical agreement. The RCs had one of the three following structures:

1. Modals: The pianist [that *will/might/could/should* lose]…

2. Have-auxiliary: The musician [that *had performed*]…

3. Past-tense + modifier: The guests [that *arrived yesterday*]…

The RCs did not contain any other NPs which potentially could have confused participants about the number of the verb because of agreement attraction (Wagers, Lau & Phillips, 2009, p. 207).

32 test items were created (see appendix IV), as well as 52 filler items (see appendix V). Fillers tested participants' sensitivity to 6 other grammatical distinctions: appropriate use of the modifiers *many* vs. *much*, selection restrictions, the count vs. mass distinction, reflexive agreement, fronted objects, and filler-gap resolution. These filler items were created with two conditions (acceptable vs. unacceptable). Fillers were included in the study in order to keep the participant from detecting the purpose of the study, namely subject-verb agreement, and to provide a baseline measurement of morphological and syntactic accuracy.

3.3.2.2 Confidence rating

Following each item, a prompt elicited participants' confidence in their answer on a 1-3 point Likert-type scale, ranging from not confident to very confident. This rating provide additional information about participants' knowledge of sentence form and structure, and logically, a correlation between accuracy and confidence is expected (Spinner & Gass, 2019, p. 77). The confidence rating was not used in later analysis, so I will not discuss it further.

3.4 Predictions of different hypotheses

In section 2.3, I discussed three hypotheses concerning L2 learners' subject-verb agreement errors: the Representational Deficit Hypothesis, the Missing Surface Inflection Hypothesis and the Functional Reassembly Hypothesis. In this section I lay out the predictions that each of the hypotheses would make about my experimental results or where participants are expected to make errors/have more difficulty.

3.4.1 Hypothesis 1 – Representational Deficit Hypothesis (RDH)

Since RDH argues that morphosyntactic features associated with agreement are only accessible to L2 learners if they have the same morphosyntactic features in their L1, the expectation is that participants will have a hard time with inflection, both with lexical verbs and auxiliaries. This might lead L2 learners to use agreement markers randomly or use default forms. This is, however, not consistent with fact and the question is thus how this hypothesis can be modified.

The assumption is that if a Norwegian speaker of L2 English produces correct agreement, this is due to superficial imitation of native speakers. Slabakova (2016, p. 187) writes that learners can imitate correct performance, which evidently means that more exposure to the target language is equivalent with more exposure to the morphosyntactic features represented in the target language – i.e. the second language learners are in a better equipped position to imitate correct agreement. As such, an overall predictions in terms of RDH *might* suggest participants in this experiment correctly will accept and reject items (9a)-(9d) more often than they correctly will accept and reject items (9e)-(9h) since auxiliaries are more common in the inflectional paradigm. RDH might also predict different outcomes for items that include main verbs. Since non-finite, bare forms are more common than inflected forms, participants in this study are expected to say "yes" more often to items (9f) and (9h), compared to items (9e) and (9g) since these items include the inflectional agreement marker *-s*.

As for auxiliaries, RDH does not make any clear predictions about which items participants will make more or less error with. A possible prediction is that participants will prefer to use one type of auxiliary (is or are) as a default form – e.g. participants might prefer sentences with "are" over sentences with "is" regardless of the grammaticality of the sentence.

3.4.2 Hypothesis 2 – Missing Surface Inflection Hypothesis (MSIH)

I will first lay out predictions regarding items including lexical verbs and then move on to predictions on auxiliaries.

MSIH is primarily able to explain errors where an affix like *-s* is missing, since this is where we are "missing" surface inflection. As such, MSIH predicts that participants in this experiment may make errors with item (9f) where surface inflection is missing. Given that MSIH only predicts errors where we are missing surface inflection, it will not make any clear predictions about item (9h) where surface inflection is not missing. In fact, MSIH predicts participants to perform accurately on all other conditions except (9f). As such, they will accept (9h) at high rates and reject (9g). The reason for predicting that participants will reject (9g) is that fact MSIH argues that second language learners have the correct grammar internalized. Based on this, participants will be good at rejecting (9g), because it violates a grammatical pattern.

All of these predictions suggest that non-finite or bare stems are used instead of finite forms – i.e., a main verb bearing the infinitival marker is not non-finite, but rather used as a default form which exhibits properties of finite verbs.

With auxiliary verbs, we are always mapping into inflected, suppletive forms regardless of whether the subject is singular or plural. This makes predictions of error rates of sentence (9a)-(9d) compared to sentence (9e)-(9h) harder in regards to the MSIH. However, as seen, Prevost & White (2000, p. 101) argue that learner sometimes use "default" nonfinite forms instead of finite forms, and that these nonfinite forms are inserted into a node which bear the [+finite] feature. If this is the case, the question is what nonfinite forms Norwegian would use for auxiliary *be.*

A possible explanation is that "are" is the default Norwegians insert into the node with the [+finite] feature. This would predict participants to use "are" in sentences where "is" is required, as we see in (9a). As such MSIH would predict increased error in (9b). Another prediction based on the default "are", is that we might expect increased error in (9c). However, since MSIH only explains errors where surface morphology is missing, this account might predict that participants should accept grammatical items with auxiliaries. Nonetheless, it is still difficult to identify what the account predicts for auxiliary conditions.

3.4.3 Hypothesis 3 – Feature Reassembly Hypothesis

FRH argues that errors in L2 are explained by challenges with reconfigurations of features from L1 to L2. As seen in chapter 2.3.2.3 and illustrated by the examples in (6), Norwegian marks present tense with *-er* in sentences with both singular and plural subjects. In English, however, present tense subjects need to agree with their verbs and are thus marked differently – i.e. Norwegian and English mark the same feature with different lexical items. It is thus reasonable to expect a possible transfer of the Norwegian morpheme *-er* to the English morpheme *-s*, leading Norwegians to overuse the morpheme *-s* (Garshol, 2019). This might lead participants in this experiment to say "no" more often to items (9f) and (9h), and "yes" to items (9e) and (9g) since these items include the *-s* morpheme. This evidently leads the participants to make errors with items (9h) and (9g).

Items including auxiliaries can be more challenging to make predictions about since the Norwegian inflectional paradigm of the verb "være" is quite different then the English inflectional paradigm of the auxiliary "be". A difference between Norwegian and English is the fact that Norwegians do not use *be* (være) in present tense like English does (see 10). Since Norwegian speakers have to inflect in both singular and plural cases in English, a

possible prediction might be that error rates could potentially be higher for auxiliary verb conditions than main verb condition.

(10)

      a. The guest that arrived **is** complaining

      b. The guests that arrived **are** complaining

      c. Gjesten som ankom klag**er**

      d. Gjestene som ankom klag**er**

The notion of use of a default form is also potentially relevant here, as seen in previous sections. Garshol (2019, p. 88) argues for the likelihood that L2 speakers will overuse more frequent forms of *be,* and seeing that "is" almost is more than twice as frequent than "are" (Davies, 2004), an expected outcome is that the auxiliary "is" is used as a default form.

3.4.4. Hypotheses summarized

The hypotheses presented above all seem to predict some problems with functional morphology for Norwegian speakers of L2 English in terms of acceptability of SVA. There are however differences in predictions of which of the eight conditions from the AJT that will be problematic for the participants. RDH believes that syntactic features that are absent from the L1 will not be acquirable, leaving L2 speakers at a non-native-like stage in their L2. Thus, since non-finite bare forms of verbs are more frequent and does not require inflection, RDH predicts participants to make fewer errors in these cases than other conditions. MSIH, on the other hand, seems to argue that participants should perform well with all conditions except for those where agreement is necessary. Since the link between morphology and syntactic knowledge is believed to be weak, MSIH argue that Norwegian participants will have problems with retrieval of the correct lexical items (e.g. 3rd person singular -s). The difference between RHD and MSIH is thus quite noticeable in the fact that RDH predicts a generally low accuracy rate while MSIH predicts a higher accuracy rate in the AJT. FRH, however, links the 3rd person singular -*s* with the Norwegian marker of present tense in both singular and plural (-*er)*, and proposes that Norwegians might overuse -*s*. This will evidently predict participants to have a high error rate for main verb items with plural subject NPs. As for auxiliary verb items, the different hypotheses all might predict the use of default forms.

Chapter 4: Results

This chapter presents the results from the AJT. Group-level results across all conditions (see section 4.1) are presented first. Statistical analysis of the group results is presented in section 4.2. Finally, individual difference results are presented in section 4.3. The aim of this study was to assess intuitions Norwegian speakers of L2 English had concerning English subject-verb agreement. The conditions in the experiment were presented in section 3.3, but are repeated here for convenience:

Condition 1: [Sg Subj – Aux – G] – *"The boy that had fallen is yelling"*

Condition 2: [Sg Subj – Aux – UnG] – *"The boy that had fallen are yelling"*

Condition 3: [Pl Subj – Aux – UnG] – *"The boys that had fallen is yelling"*

Condition 4: [Pl Subj – Aux – G] – *"The boys that had fallen are yelling"*

Condition 5: [Sg Subj – Main – G] – *"The boy that had fallen yells loudly"*

Condition 6: [Sg Subj – Main – UnG] – *"The boy that had fallen yell loudly"*

Condition 7: [Pl Subj – Main – UnG] – *"The boys that had fallen yells loudly "*

Condition 8: [Pl Subj – main – G] – *"The boys that had fallen yell loudly"*

4.1 Group-level results

Figure 1 shows the overall error rates for the target items in percentages (blue columns represent grammatical conditions, while orange represent ungrammatical conditions), while Table 1 shows the mean (and standard deviation) of the raw scores for each test condition in the AJT. As presented in section 3, participants were asked to rate the test items as good or bad. These answers were later coded, such that the answer "*good"* was given value 1 and *"bad"* was given the value 0. Any value over .5 can thus be considered a condition where participant performance was above chance.
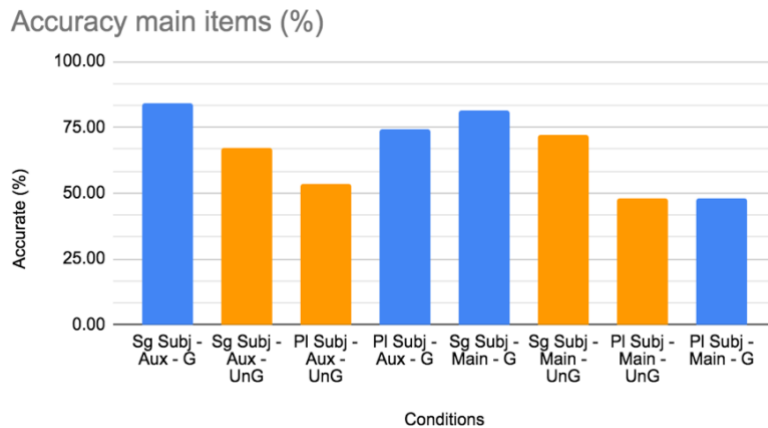
Figure 4: Average accuracy for test items.

	Sg Subj- Aux	Pl Subj – Aux	Sg Subj – Main	Pl Subj – Main
GRAM	0.84 (.37)	0.74 (.44)	0.81 (.39)	0.48 (.50)
UNGRAM	0.67 (.47)	0.53 (.50)	0.73 (.45)	0.53 (.50)

Table 3: Mean scores (Standard Deviations) for test items.

As Table 3 demonstrates, some conditions can be interpreted as near chance performances, which is also visible in Figure 1. The figure and table show that condition "Pl Subj-Main-G" (M = 0.48, SD = .50), "Pl Subj-Main-UnG" (M = 0,53, SD = .50) and "Pl Subj-Aux-UnG (M = 0.53, SD = .50) can be interpreted as conditions where participants performed near chance since the mean score is close to .5. However, the table shows that the remaining conditions are conditions where the average score is higher than chance, i.e. the participants have low error rates for both grammatical conditions and ungrammatical conditions since the mean scores are above .5.

The mean scores for the ungrammatical items are overall lower than the score for grammatical items. Norwegian speakers of L2 English have a higher error rate with ungrammatical sentence than grammatical sentences– i.e. L2 speakers have a tendency to say "yes" to ungrammatical sentences more often than "no" to grammatical forms.

By comparing items containing plural subjects with items containing singular subjects, it is evident that the participants in the study have a higher error rate with plural subject items since their overall mean scores are lower than for items with singular subjects. The only exception is condition "Pl Subj-Aux-G" (M = 0.74, SD = .44).

The most striking finding is the fact the participants are at chance when judging whether conditions with main verbs and plural subjects (Pl Subj-Main) are grammatical or

28

not. We see that in both "Pl Subj-Main-G" and "Pl Subj-Main-UnG" conditions, accuracy is at chance. This suggests that participants, on the whole, have significant difficulty deciding on the appropriate agreement for a main verb when the subject is plural.

A final interesting aspect to draw from Figure 4 and Table 3 is the distinction between auxiliaries versus main verbs. As the table demonstrates, we see a tendency of auxiliaries having a higher accuracy rate than main verbs.

	Grammatical	Ungrammatical
Mean	0.76 (.40)	0.80 (.30)

Table 4: Filler items categorized into grammatical and ungrammatical items.

Table 4 summarizes the mean- and standard deviation values for the filler items. The table shows that all grammatical filler items have been accepted as good sentences, while the ungrammatical filler items have been rejected by the L2 participants. In comparison to the target sentences, we see that ungrammatical filler sentences overall had a lower error rate than ungrammatical target sentences, which was also seen for grammatical filler – and grammatical main items. These results portray the fact that participants generally had an overall high accuracy score and understood the task.

4.2 Statistical analysis [3]

Above, the results and differences between the conditions were described informally. In order to determine which differences and effects were statistically reliable, a statistical analysis was conducted.

Accuracy scores were analyzed using logistic mixed-effects regression in R (R Core Team, 2020) using the lme4 (Bates, Mächler, Bolker & Walker 2015) and lmerTest (Kuznetsova, Brockhoff & Christensen, 2017) packages. Logistic regression was used instead of a standard ANOVA because we used binary response proportion (0 v. 1) rather than a continuous variable like reaction time. The model included Verb Type (main v. auxiliary), Subject Number (singular v. plural), Grammaticality (grammatical v. ungrammatical) and their interaction as contrast-coded fixed effects. The model also contained random intercepts for participant and item.

---

[3] Assistance with running the models was received from Dave Kush. I informed him of what contrasts I wanted and what comparisons to make. The results and the comparisons were then discussed.

4.2.1 Results

A summary of statistical effects from the model is presented in Table 5:

	Coefficient (SD)	z-value	p-value
Intercept	0.33 (0.14)	2.354	0.019
VerbType	-0.62 (0.20)	-3.032	0.002
SubjectNumber	0.93 (0.15)	6.034	<.001
Grammaticality	0.38 (0.20)	1.906	0.057
VType:SubjNum[4]	0.65 (0.31)	2.104	0.035
VType:Gramm	-1.15 (0.40)	-2.844	0.004
SubjNum:Gramm	0.37 (0.31)	1.210	0.226
VType:SubjNum:Gramm	0.68 (0.62)	1.116	0.264

Table 5: summary of statistical effects.

When looking at the statistical effects from the model, we found both main effects and interaction effects. A main effect is when one factor has an overall average effect independent of other factors. An interaction effect, however, is when the combinations of two effects seem to matter, like we see in the last four rows in Table 5.

According to the model there were two significant main effects: A main effect of Verb Type reflected the fact that conditions with auxiliary verbs were rated more accurately on average than conditions with main verbs ($p < .01$). A main effect of Subject Number reflected the fact that conditions with singular subjects were rated more accurately on average than conditions with plural subjects ($p < .001$). Finally, there was a marginally significant main effect of Grammaticality, which reflected that grammatical items were rated more accurately on average than ungrammatical items. These effects align with the informal descriptions in section 4.1

The main effects above are qualified by two significant interaction effects: First, Verb Type interacted with Subject Number ($p < .05$). This reflects the fact that in conditions where the subject was singular, the average accuracy was roughly equal when the verb was a main or auxiliary verb (roughly 76%). However, when the subject was plural, accuracy was higher on average in auxiliary verb conditions (64%) compared to main verb conditions (50%).

---

[4] "VType:SubjNum" means "Interaction of Verb Type and Subject Number". All ":" in this table refer to "interaction".

30

Second, Verb Type interacted with Grammaticality ($p < .01$). In ungrammatical conditions accuracy did not differ depending on whether the verb was a main verb or an auxiliary (roughly 61%). However, in grammatical conditions overall accuracy was affected by verb type: average accuracy was higher when the verb was an auxiliary (79%) than when it was a main verb (65%). No other effects achieved significance in the model.

The effects from the logistic regression analysis conform to the descriptions made in the informal analysis of the group-level results in regards to main effects. The informal description of the results found that participants on average had higher error rates for ungrammatical items than grammatical items, which is in line with the effect seen in the model (Grammaticality), displayed by p-value 0,057. This effect was, however, only marginally significant. The other effect found in the informal description was participants' error rates for items containing singular and plural subject NPs – participants had higher error rates for items with plural subjects, they were in fact at chance when judging these conditions. This also conforms to the main effect (SubjectNumber, p-value = <.001) found in the logistic regression analysis. A final main effect from the informal description was participants accuracy scores for items with main verbs and auxiliaries – participants had higher accuracy rates in items with auxiliary verbs than items with main verbs, which again is seen in the model by the main effect VerbType (p-value (.019).

However, what was not found in the informal analysis of the group-level results was the interaction effects. Thus, the informal analysis of the group-level results only conformed with the main effects found in the model.

4.2.2 Individual Difference Results

I was also interested in knowing how accuracy varied across individuals and how accuracy was affected by exposure to English. This was an interesting aspect to consider since there exists different views as to how L2 input affects L2 acquisition and thus the acquisition of English SVA. Research has shown that SVA errors occur well into advanced stages in L2 acquisition (e.g. Garshol, 2019) which could be indicative of L2 input not affecting L2 accuracy to great degree. However, different studies (e.g. Mañoz, 2014) state that cumulative exposure to the target language and contact with high-quality input is a strong indication for good outcomes in the target language (Mañoz, 2014, p. 463). This is also seen in Slabakova (2016, p. 142), which states that type of exposure and length of target language exposure may matter equally much or more than age of convergence. Slabakova also notes that the Representational Deficit Hypothesis assumes that correct subject-verb agreement production

by Norwegian speakers of L2 English is due to superficial imitation of native speakers (2016, p. 187), which again might indicate the importance of L2 input in acquisition.

Based on these findings, I wanted to investigate individual differences since I hypothesized that there would be a correlating relationship between how many hours the participants were exposed to English per week and accuracy. I chose English Media Scores as my independent variable as this variable provides a valid indication of how much of the target language the participants were exposed to on a weekly basis. I also wanted the individual analysis model to look at differences between main verbs and auxiliaries since none of the hypotheses presented in chapter 3.4 gave clear predictions for how accuracy rates for auxiliary verbs would do in the AJT. Individual differences could possibly present some important aspects into the auxiliary verb discussion.

4.2.2.1 Statistical Analysis

Individual participant accuracy scores were analyzed using a linear mixed effects model implemented in R. The model used Verb Type (main v. auxiliary), Grammaticality (grammatical v. ungrammatical), participants' self-reported English Media score, and their interactions as fixed effects. The model contained random intercepts for participant.

Figure 5 plots individual participant accuracy by the number of hours of English media each participant reported to consume in a typical week.[5] The figure also separates accuracy by grammaticality and verb type in order to visualize whether accuracy varied substantially for different item types.

---

[5] Dave Kush created Figure 5 in R.
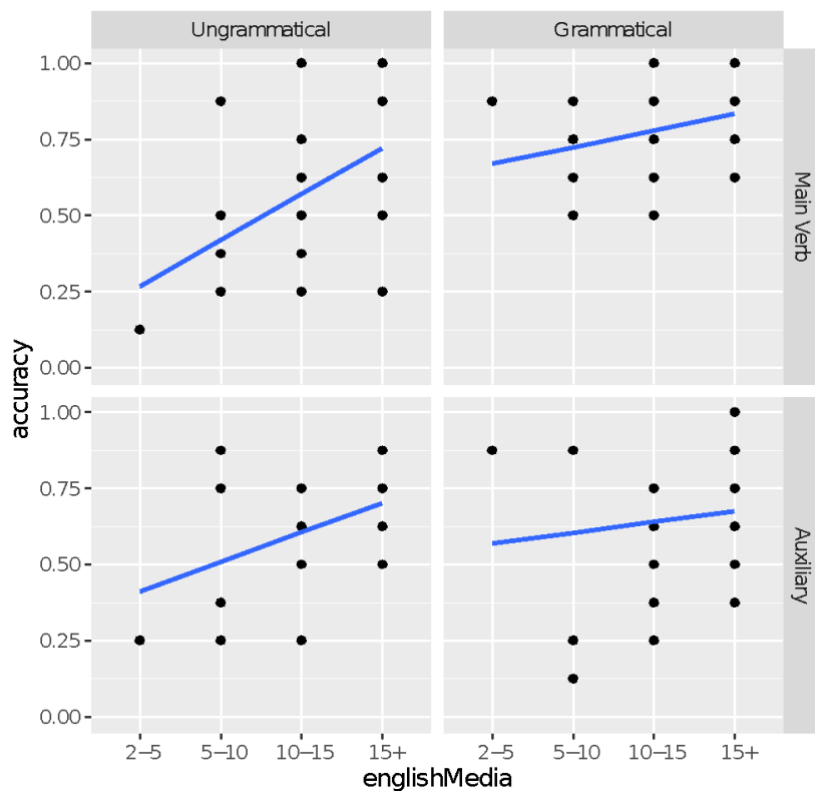
Figure 5: Individual participant accuracy.

The model revealed a main effect of grammaticality: participants were more accurate on average on grammatical sentences than ungrammatical sentences ($t = 3.49, p < .001$). There was also a main effect of English Media exposure: participants' accuracies were on average higher the more exposure to English Media they reported. This can be seen in the fact that the lines in Figure 5 have a positive slope. This was in line with what was hypothesized for the individual differences analysis – exposure to the target language has a positive correlation with accuracy for the participants.

There was also an interaction between grammaticality and English Media exposure: English Media had a larger effect on accuracy in ungrammatical conditions than in grammatical conditions ($t = 2.89, p < .01$). There were no other main effects or interactions. Based on these findings we can argue that L2 Norwegians are getting better at identifying ungrammatical English sentences with more exposure to English Media.

Chapter 5: Discussion

The first part of the discussion gives a brief summary of the relevant effects (see chapter 4) and presents some main findings before interpreting how the results align with the predictions of the three hypotheses discussed in Chapter 3.3. The individual scores are discussed in section 5.4. The final part of the discussion lays forth strengths and weaknesses of the study and presents some relevant suggestions for further research.

5.1 Main findings from the AJT

This study set out to investigate intuitions L2 Norwegian speakers had about subject-verb agreement constructions and three different hypotheses were presented in order to make predictions as to how Norwegians would perform (see section 2.3.2-2.3.4 and section 3.3). As presented in chapter 4, the results revealed three main effects: (1) the average accuracy was marginally higher in grammatical conditions than in ungrammatical conditions, (2) error rates were higher overall when the subject NP was plural ($p < .001$) and (3) items with auxiliary verbs were rated more accurately on average than items with main verbs ($p < .01$).

In addition to this, the statistical analysis revealed interaction effects in two different cases. First, an interaction was seen between Verb Type and Subject Number. In conditions where the subject was singular, the average accuracy was roughly equal when the verb was an auxiliary or a main verb. However, when the subject was plural, accuracy was higher on average in auxiliary verb conditions. This interaction was driven by the fact that participants had low accuracy when picking the correct agreement with a plural subject and a main verb. That is, participants were at chance in condition "Pl subj-Main-Ung" and "Pl subj-Main-G". Secondly, an interaction between Verb Type and Grammaticality was found. In grammatical conditions, average accuracy was higher when the verb was an auxiliary than a main verb, while for ungrammatical conditions the Verb Type did not influence accuracy.

For all the three hypotheses it was challenging to make explicit predictions for auxiliary verbs, which in large was due to the fact that English auxiliaries always map to inflected, suppletive forms regardless of whether the subject is singular or plural. However, based on all the three hypotheses it might be argued for the use of default forms – i.e. that Norwegian speakers prefer either "is" or "are". When looking at the interaction effects from the statistical analysis we see that when the subject NP was plural, accuracy was higher on average in auxiliary verb conditions. This might indicate that Norwegian speakers in general might use "are" as a default form

Considering the accuracy scores from the AJT, it is evident that the participants in this study show a tendency of struggling, in general, with English subject-verb agreement. This was an expected effect and in accordance with SLA research in general, where it has been observed that L2 learners are inconsistent in their use of correct morphology in their target language (White, 2003, p. 178). As seen in the theoretical backdrop for this thesis, many explanations have been offered in order to grasp why people in general might have problems with agreement in English.

Work in SLA has hypothesized and showed that the native language of a learner will influence the outcome of the L2 language, and that an already-internalized grammar is transferred to or interferes with L2 acquisition (White, 2003; Shimanskaya, 2015; Ortega, 2013). As such, it was plausible to assume that Norwegian speakers would have problems with English subject-verb agreement seeing as how Norwegian does not have subject-verb agreement. Thus, by looking at the results from the AJT, I interpret this to mean that there is a general problem with agreement independent of the three hypotheses presented in section 3.3.

I now proceed to discuss the main effects in addition to the interactions in light of the three hypotheses. The first main effect stating that accuracy was higher with grammatical conditions than ungrammatical conditions will only be addressed shortly since this effect in large degree conform to SLA research in general (e.g. Jensen, Slabakova, Westergaard, 2017, p. 339; Jensen, Slabakova, Westergaard, Lundquist, 2020). Research done by Jensen et. al. found that participants overall have a preference for grammatical sentences over ungrammatical sentences (2019, p. 18). The same was found in Jensen et. al (2017, p. 341) where participants made fewer errors with grammatical than ungrammatical sentences, which is in line with what was found in the AJT results in the present study. Grammatical sentences are arguably easier to judge correctly than ungrammatical sentences because we more often are exposed to correct grammatical forms.

The remaining main effects and interaction, on the other hand, will be discussed in light of the three hypotheses.

5.2 AJT results and Representational Deficit Hypothesis

RDH is concerned with the underlying knowledge of syntactic representation and holds that poor morpho-syntactic performance in L2 is due to incomplete grammatical knowledge (Slabakova, 2016). The RDH predicts that Norwegian participants in this study should not have person and number features on T in their English L2 grammar because Norwegian T

does not have person and number agreement features (e.g., Harwood, 2014; Carnie, 2013; Agder 2003), and based on this RDH might be able to explain some of the effects from the AJT: The RDH predicts that Norwegian speakers will, in general, struggle with English agreement if English agreement features are not available to Norwegian speakers. We see that Norwegian speakers of L2 English do, in fact, make a number of agreement errors.

However, there are also cases where RDH falls short in its predictions. First, we saw RDH predicting that participants would say yes more often to items containing bare forms (i.e. item "Sg subj-main-UnG" and "Pl subj-main-G") since these structures are more common and Norwegian L2 speakers thus use these as "default" forms. However, we saw that participants' accuracy scores were markedly different for items with plural and singular subject NPs: participants rated items with singular subject NPs – where overt agreement morphology is necessary and the correct verbs are not in bare/"default" form – more accurately on average than conditions with plural subject NPs, where bare/"default" forms are grammatical ($p < .001$) (e.g. in line with Jensen et. al., 2017, p. 342). This is problematic for RDH, since RDH expected participants to have problems with cases where overt agreement is necessary (which is not the case for item "Pl subj-main-G"). This could be an indication that it is in fact not the case that poor morpho-syntactic performance in L2 is due to incomplete grammars, since the L2 participants in this study seem to perform with a certain accuracy both for cases where finite and non-finite bare forms are present. The results also revealed that participants essentially were at chance when judging whether main verb conditions with plural subject NPs were grammatical or not, which might contradict RDHs predictions that participants would have high accuracy scores for items "Pl Subj-main-G", where they should use bare forms.

We also saw that participants' accuracy scores for auxiliaries were higher than for main verbs. It is not clear if this effect is predicted by RDH: if agreement on auxiliary verbs is controlled by the same number and gender features as with main verbs, why should auxiliaries be any better? One way that the RDH might try to accommodate the difference between main and auxiliary verbs is as follows: RDH argues that more exposure to the target language results in L2 speakers developing strategies to 'fake' agreement in specific contexts. Since auxiliaries are so frequent, one might assume that participants have developed more reliable strategies to 'fake' proper agreement with auxiliaries than main verbs. In addition to this, Norwegian schools put great emphasis on teaching the inflectional paradigm of auxiliaries (Jensen, et. al., 2017, p. 345; Jensen, et. al., 2020, p. 22), which might cause

Norwegian speakers of L2 English to be more aware of this inflectional paradigm than the main verb inflectional paradigm.

5.3 AJT results and Missing Surface Inflection Hypothesis

MSIH proposes that L2 speakers in general struggle with retrieval of the lexical items necessary for successful inflection (3rd person sg, *-s*) (Slabakova, 2916; Prevost & White, 2000). Since MSIH proposes that the problem lies in retrieving the overt morpheme *-s*, its only direct prediction is that participants should struggle with main verbs where *-s* was necessary (e.g. condition "Sg Subj-main-G").

The first main effect (error rates for items with plural v. singular subject NPs), can be argued to be a mark against MSIH. Since MSIH proposes that L2 learners have acquired the relevant features of a syntactic node and that the impairment lies in the retrieval of the overt morphological form, conditions with plural subject NPs would be predicted to have a higher accuracy score than conditions with singular subject NPs.

The largest problem for MSIH is the finding that that participants were at chance with accuracy in both conditions with main verbs and plural subject NPs (grammatical and ungrammatical). This means that participants accepted ungrammatical agreement like *"*the guests complain-s"* (Pl Subj-main-UnG) close to 50% of the time and that they rejected grammatical agreement like *"the guests complain"* (Pl Subj-main-G) just as often. MSIH does not explain why participants accept the sentences in the ungrammatical "Pl Subj-main-UnG". Why should people mistakenly retrieve the *-s* with a plural subject? MSIH also does not explain why people should reject "Pl Subj-main-G", where there is not a need for any overt inflection.

For the final main effect (auxiliaries having a higher accuracy rate than main verbs on average), it was difficult to determine if this was compatible or incompatible with MSIH. MSIH seems to only explain errors where surface inflection is missing and a 'bare' form is chosen. There is, however, no 'bare' form of the auxiliary since it is suppletive. As discussed in section 5.1, the notion of a 'default' form might be a possible explanation for accuracy errors with auxiliaries. Prevost and White (2000) argue that learners sometimes insert default nonfinite forms into a node which bears the [+finite] feature. Inserting the non-finite "be" would not be the correct default form for English auxiliaries, so a different 'default' form would need to be used. The question is thus which default forms Norwegians would use. One possibility is that "are" is used as a default form. If "are" is the default form that is substituted

when retrieval fails, MSIH predicts that participants would have more problems with the condition "Sg Subj-aux-G" where "is" should be picked. MISH might also predict participants to accept "Sg Subj-aux-UnG, where "are" is used instead of "is". Based on the results from the AJT and seen in the section above, accuracy was higher on average in auxiliary verbs when the subject NP was plural, which again might be seen as evidence for "are" as a default form since the speakers seem to accept conditions with "are", and incorrectly reject grammatical conditions and accepts ungrammatical conditions with "is".

In sum, even though performance with auxiliaries might be explained, the results from the AJT seem to give an overall mark against MSIH since we see that participants in general have problems with subject-verb agreement. According to MSIH participants should in theory only struggle with singular subjects where agreement is needed, which is in contrast to what was found in the AJT.

5.4 AJT results and Feature Reassembly Hypothesis

Since FRH argues that L2 impairment lies in the organization and reassembly of features from a speakers L1 to their L2, an issue with subject-verb agreement was raised by Garshol (2019) where he questions whether Norwegian L2 speakers equate the Norwegian present tense marker -*er* with the English inflectional morpheme -*s*. With this assumption in mind, FRH predicted participants in the study to say *yes* more often to main verb conditions that included the -*s* morpheme, and naturally *no* to the other main verb conditions, which evidently causes participants to make errors in conditions with plural subject NPs.

The fact that participants seem to have high error rates with main verbs and plural subjects can thus be seen in line with FRH. If, as proposed by Garshol (2019), L2 Norwegian speakers want to use -*s* as a basic marker for "present", then the participants would want to use this marker with both singular and plural items. As such, FRH would expect participants to say *yes* to sentences that include both plural subject NPs *and* the inflectional marker -*s* (e.g. "Pl Subj-main-Ung"), which is what we see from the results. These results can thus be seen in line with the prediction that Norwegians equate the present tense marker -*er* with the English inflectional morpheme -*s*. Based on this finding, it is plausible to assume feature reassembly to be a problematic part in second language acquisition. It seems that Norwegian speakers have not fully managed to implement the necessary modifications to person and number features in English. As such, they accept sentences where the -*s* morpheme is present, which

might reflect an association to the inflectional morpheme *-er* which is associated with [+pres] in Norwegian.

Auxiliaries are, as indicated in previous sections, difficult in regards to making correct predictions since we always have to inflect in both singular and plural cases. However, an interesting contribution to the discussion of auxiliary verbs and correct feature reassembly has been presented by Garshol (2019, p. 88). He argues that since auxiliary *"BE"* has suppletive forms in all finite contexts, it would be plausible to assume that Norwegian L2 learners would only use *one* of the suppletive BE forms for all persons, as discussed previously. Moreover, it is reasonable to assume that the L2 learners would use the most frequent form of *BE* - in this study *"is"*, which is more than twice as frequent as *"are"* according to the British National Corpus (Davies, 2004). As seen, the results from the AJT and presented above, "are" might be predicted as a default form, which is not in line with the prediction put forth by FRH. It thus seems as the feature reassembly hypothesis alone cannot fully explain the error pattern or the success pattern for suppletive agreement among the Norwegian participants.

5.5 Individual scores

In addition to looking at group results from the AJT, individual participant accuracy scores were also analyzed (see section 4.2.2). The model tested for effects Verb Type, Grammaticality, participants' self-reported English Media exposure and their interactions.

The most interesting finding from the individual difference analysis was that an interaction between Grammaticality and English Media exposure was present. What this interaction shows is that the more exposure the participants have to English Media, the more accurate they are in rejecting ungrammatical items. English Media exposure does not correlate as strongly with accuracy in grammatical cases: the slope of the line for grammatical items in Figure 5 is almost entirely flat. As such, whether or not participants have a lot of exposure to English Media does not predict the accuracy for grammatical cases. It is difficult to explain exactly why the effect is only present with ungrammatical items and not grammatical items. However, if participants make incorrect judgements for ungrammatical items, they are picking out the wrong verb for agreement. Thus, the results are stating that Norwegians' intuitions about ungrammatical forms are getting better with more exposure to English Media. This might be seen in line with any of the three hypotheses. RDH predicts that more exposure to the target language implies stronger "faking" rules. MSIH predicts that

more exposure equals stronger links between T heads and their morphological exponents. FRH predicts that more exposure would mean more evidence for feature reassembly.[6]

5.6 Strengths and weaknesses

Summing up the arguments made so far, we see that some of the effects from the AJT can be explained by the three hypotheses considered while other effects are harder to explain. No single hypothesis can easily predict the full range of effects that we observed. In light of this, it is important to note that the hypotheses address different aspects of L2 acquisition – e.g. MSIH is a *performance* based hypothesis while RDH and FRH are more concerned with *competence*. It is thus reasonable to assume that the different hypotheses explain different aspects of the effects seen from the AJT. It might in fact be the case that some morphemes *are* harder to retrieve than others (as MSIH predicts), but that retrieval of certain morphemes is not the only problem L2 speakers have with acquisition. It is also important to note that the three hypotheses in theory cannot be true simultaneously, since their assumptions of L2 acquisition is in conflict with each other; RDH holds that L2 grammar is fundamentally impoverished (i.e. Norwegian L2 speakers of English do not have features in T), while MSIH and FRH argue that L2 grammar has all of the features that L1 English has in T. From the results, it seems the real challenge is acquiring English agreement structures in general that is not present in the participants' own native language.

Things to have in mind are methodological limitations of the study and the question of whether these results can be generalized. The main limitation of the study is the fact that a control group was not recruited. As stated in chapter 3.2 an Acceptability Judgement Task was planned for native speakers. However, since recruiting enough participants was problematic, the control group was not included in this study. A control group of native speakers should preferably have been included in order to see what "native-like" performance would have been in regards to the eight conditions in the AJT. For example, if the control group showed a high error rate in general for one item, this could imply that the item should be discarded and the study possibly redesigned (Spinner & Patti, 2019, p. 86). Since the study did not include a control group, these factors were not controlled for.

---

[6] Feature reassembly, due to more exposure, is in line with the notion of *restructuring* under the Full Transfer/Full Access hypothesis (Schwartz and Sprouse, 1996). The model claims that the initial state of L2 acquisition is equal to the final state of L1 acquisition, but as learners gain positive evidence from L2 input, the initial state will change and the learners will be able to restructure their L2 grammar.

Further, a binary yes/no format was chosen for the AJT which also could potentially create a few challenges (see section 3.1), one being the issue of grammaticality versus acceptability. In an AJT, participants are asked to rate the acceptability of a sentence and seeing that acceptability is equal to percepts and that there is no direct way of measuring percepts, participants are asked report their perceived acceptability on a given scale (Schütze & Sprouse, 2012, p. 3). By making this "scale" a yes/no option, it forces the participants to make clear judgements of whether they think sentences are acceptable or not, which might be problematic seeing that acceptability might not be as straight forward, and potentially lead to biased answers (see section 3.1.1).

5.7 Suggestions for further research

There are several possibilities for further investigation which could shed more light on L2 acquisition of subject-verb agreement. I would argue that the study accentuates the need for more research since the main finding seems to portray participants inability to make correct agreement judgements in general.

First, the present study was formatted in a way that would allow Signal Detection Theory (SDT) to be used for analyzing the results. SDT is an analytical tool which has been used in perceptual studies and then been adopted as an analytical theory for acceptability judgements specifically (Huang & Ferreira, 2020). Huang & Ferreira (2020) write that "STD assumes that performance is not perfect and describes how well observers can discriminate or recognize certain signals given background noise". Thus, if the study implemented the use of SDT for analyzing the data in this study, it would have allowed us to quantify bias and sensitivity in a way that only comparing the average accuracy did not.

Further, much research on subject-verb agreement has been corpus-based (e.g. Garshol, 2019; Breiteneder, 2005), while the present study implemented an AJT to investigate the matter. It would thus be interesting to see if more experimental research with specific emphasis on production would provide new insight into acquisition of subject-verb agreement. Acceptability judgements do not provide us with direct access into linguistic competence, but allows us to make inferences about linguistic knowledge. Production-based studies could provide useful information about subject-verb agreement from a more performance-based aspect which the present study does not provide.

Chapter 6: Conclusion

This study set out to investigate L1 Norwegian speakers´ intuitions about English subject-verb agreement, which is known to be a problematic area for L2 speakers of English. Previous research has found that L2 learners in general struggle with acquisition of this English feature, which is in line with the overall results from this study.

This study conducted an Acceptability Judgement Test in order to investigate the issue. The AJT design manipulated subject-verb constructions over three factors (Grammaticality, Verb Type and Subject Number), to test the Representational Deficit Hypothesis, Missing Surface Inflection Hypothesis and Feature Reassembly Hypothesis as explanations for why acquisition and use of morphosyntactic features might be difficult for L2 speakers.

The results from the AJT found three main effects stating that participants had higher accuracy rates for grammatical than ungrammatical items, participants accuracy scores were higher when the subject NP was singular than plural, and finally, participants' accuracy rates were higher with auxiliaries than main verbs. The results also revealed interaction effects, showing that participants were less accurate with plural agreement when the verb was a main verbs than an auxiliary, and that errors with auxiliary verbs were rejected more consistently than errors with main verbs. In addition to this, individual scores showed that the L2 speakers performed better in the AJT with more exposure to the target language.

When looking at the results in light of the predictions made by the three hypotheses above, it is evident that none of the hypotheses fully can explain the results. Since the hypotheses looked at different aspects of L2 acquisition, they attested to different parts of the results. It is also the case that the hypotheses cannot simultaneously be true given that the hypotheses assume radically different things about the nature of L2 grammar. MSIH, which is a performance-based hypothesis, was rendered as a mark against the results since it predicted high accuracy rates for all conditions except conditions with singular subjects where agreement is needed. RDH can also be seen as a mark against the results as many of RDHs predictions falls short. It is therefore arguably not the case that poor morpho-syntactic performance in L2 is due to incomplete grammars. RDH *does* expect participants' overall accuracy scores to be poor, but its predictions falls short in several other areas. Lastly, FRH seems to explain some of the effects, as participants accept sentences with the inflectional marker *-s* at high rates, which FRH argues Norwegians equate with the inflectional marker *-er,* associated with [+pres] in Norwegian.

It seems the overall challenge for Norwegian speakers of L2 English is to acquire features in the target language which are not present in their L1. The Norwegian intuitions towards English subject-verb agreement can thus be argued to be somewhat weak.

References

Acuña-Fariña, J., C. (2012) Agreement, Attraction and Architectural Opportunism. *Journal of Linguistics*, *48*(2), 257 – 295. Cambridge University Press. https://doi.org/10.1017/S0022226712000084

Adger, D. (2003). *Core syntax: A minimalist approach*. Oxford University Press.

Bailey, N., Madden, C., & Krashen, S. (1974). Is There a 'Natural Sequence' in Adult Second Language Learning? *Language Learning, 24*(2). 235-243. https://doi.org/10.1111/j.1467-1770.1974.tb00505.x

Bates D., Mächler M., Bolker B., Walker S. (2015). "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

Breiteneder, A. (2005). The naturalness of English as a European lingua franca: The case of the 'third person -s '. *Vienna English Working Papers, 14*(2), 3-26.

Carnie, A. (2013). *Syntax. A generative Introduction* (3rd ed.). Wiley-Blackwell.

Chomsky, N. (1965). *Aspects of the theory of syntax.* The M.I.T Press.

Cowart, W. (1997). *Experimental Syntax : Applying Objective Methods to Sentence Judgments*. Thousand Oaks: SAGE Publications.

Davies, M (2004-). BYU-BNC. (Based on the National Corpus from Oxford University Press.). http://corpus.byu.edu/bnc

Domínguez, L., Arche, M. J., Myles, F. (2011). Testing the predictions of the Feature Assembly Hypothesis: Evidence from the L2 acquisition of Spanish Aspect Morphology. In N. Danis, K. Mesh, & H. Sung (Eds.), *Proceedings of the 35th Annual Boston University Conference on Language Development* (pp. 183-196). Somerville, MA: Cascadilla Press.

Drummond, A. (2013). IbexFarm. Retrieved from https://spellout.net/ibexfarm/

Ellis, N. (2008). Implicit and explicit knowledge about language. In J. Cenoz & N. H. Hornberger (eds.), *Encyclopedia of Language and Education, 2nd Edition, Volume 6: Knowledge about Language.* Springer Science + Media LLC.

Ellis, R. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching.* Bristol: Multilingual Matters.

Eubank, L., Bischof, J., Huffstutler, A., Leek, P., & West, C. (1997). "Tom Eats SlowlyCooked Eggs": Thematic-Verb Raising in L2 Knowledge. *Language Acquisition, 6*(3), 171-199. https://doi.org/10.1207/s15327817la0603_1

Garshol, L. (2019). *I Just Doesn't Know. Agreement errors in English texts by Norwegian L2 learners: Causes and remedies* (Doctoral Dissertation). University of Adger.

Gibson, E. & Tomas. J. (1999). Memory limitations and structural forgetting: the perception of complex ungrammatical sentences as grammatical. *Lang. Cognit. Process, 14*, 225-248. https://doi.org/10.1080/016909699386293

Greenbaum, S., & Nelson, G. (2009). *An Introduction to English Grammar* (3rd ed.) London and New York: Routledge.

Harwood, W. (2014). Rise of the Auxiliaries: a case for auxiliary raising vs. affix lowering. *The Linguistic Review, 31*(2), 295-362. https://doi.org/10.1515/tlr-2014-0001

Holmberg, A., & Platzack, C. (1995). *The Role of Inflection in Scandinavian Syntax*. New York: Oxford University Press.

Huang, Y. & Ferreira, F. (2020). The Application of Signal Detection Theory to Acceptability Judgements. *Frontiers in Psychology, 11*(73). Doi: 10.3389/fpsyg.2020.00073

Hudson, R. (1999). Subject-verb agreement in English. *Language and Linguistics, 3*(2), 173-207. https://doi.org/10.1017/S1360674399000210

Ionin, T., & Wexler, K. (2002). Why is "is" easier than "-s"?: acquisition of tense/agreement morphology by child second language learners of English. *Second Language Research, 18*(2), 95-136. https://doi.org/10.1191/0267658302sr195oa

Jensen, I. N., Slabakova, R., Westergaard, M. (2017). The Bottleneck Hypothesis in L2 acquisition: A study of L1 Norwegian speakers' knowledge of syntax and morphology in L2 English. *Proceedings of the 41st annual Boston University Conference on Language Development,* ed. Maria LaMendola and Jennifer Scott, 333-346.

Jensen, I. N., Slabakova, R., Westergaard, M., Lundquist, B. (2020). The Bottleneck Hypothesis in L2 acquisition: L1 Norwegian learners' knowledge of syntax and morphology in L2 English. *Second language Research, 31*(1), 3-29. https://doi.org/10.1177/0267658318825067

Kuznetsova A., Brockhoff P. B., Christensen R. H .B. (2017). "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software*, *82*(13), 1–26. doi: 10.18637/jss.v082.i13.

Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research, 25*(2), 173-227. https://doi.org/10.1177/0267658308100283

Mañoz, C. (2014). Contrasting Effects of Starting Age and Input on the Oral Performance of Foreign Language Learners. *Applied Linguistics, 35*(4), 463-482. https://doi.org/10.1093/applin/amu024

Mayo, M. P. G & Olaizola, I. V. (2010). The development of suppletive and affixal tense and agreement morphemes in the L3 English of Basque-Spanish bilinguals. *Second Language Research, 27*(1), 129-149. https://doi.org/10.1177/0267658310386523

Prevost, P. & White, L. (2000). Missing Surface Inflection or Language Impairment in second language acquisition? Evidence from tense and agreement. *Second Language Acquisition 16*(2), 103-133. https://doi.org/10.1191/026765800677556046

Ortega, L. (2013). *Understanding second language acquisition* (pp. XV, 304). Routledge.

Radford, A. (2004). *Minimalist syntax: Exploring the Structure of English.* Cambridge University Press.

Ramchand, G., & Svenonius, P. (2014). Deriving the functional hierarchy. *Language Sciences 46*, 152-174. https://doi.org/10.1016/j.langsci.2014.06.013

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/

Rebuschat, P. (2013). Measuring Implicit and Explicit Knowledge in Second Language Research. *Language Learning, 63*(3), 595-626. https://doi.org/10.1111/lang.12010

Ritchie, W. and Bhatia, T. (eds.) (2009) 'Introduction to part II', The New Handbook of Second Language Acquisition. 2nd edition. Bingley: Emerald Group Publishing Limited.

Schütze, C. T., and Sprouse, J. (2012). Judgment Data. In R. J. Podesva and D. Sharma (Eds.), *Research Methods in Linguistics*. Cambridge: Cambridge University Press.

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research, 12*, 40-72. https://doi.org/10.1177/026765839601200103

Shimanskaya, E. M. (2015). *Feature reassembly of semantic and morphosyntactic pronominal features in L2 acquisition* (Doctoral dissertation). University of Iowa, Iowa. https://doi.org/10.17077/etd.p1k1k73b

Slabakova, R. (2016). *Oxford Core Linguistics. Second Language Acquisition.* Oxford University Press

Spinner, P. & Patti S. M. (2019). *Using Judgements in Second Language Acquisition Research.* Milton: Routledge.

Tremblay, A. (2005). Theoretical and Methodological Perspectives on the use of Grammaticality Judgement Tasks in Linguistic Theory. *Second Language Studies, 24*(1), 129-167.

VanPatten, B., & Benati, A. G. (2010). *Key Terms in Second Language Acquisition.* Bloomsbury Publishing PLC.

Wagers, M. W., Lau, E. F. & Phillips C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language, 61*, 206-237. https://doi.org/10.1016/j.jml.2009.04.002

White, L. (2003). *Second language acquisition and universal grammar* (pp. XV, 316). Cambridge University Press.

Wold, S. H. G. (2017). *INGlish English: The progressive construction in learner narratives* (Doctoral dissertation). University of Bergen, Bergen.

Zobl, H., & Liceras, J. (1994). Review Article: Functional Categories and Acquisition Orders. *Language Learning 44*(1), 159-180. https://doi.org/10.1111/j.1467-1770.1994.tb01452.x

Appendices

Appendix I: Pedagogical implications (relevance for the teaching profession)

Seeing that English subject-verb agreement is an area of the English language found to be particularly problematic for L2 learners, this thesis is undeniably relevant for the teaching profession. A vital part of the job of an ESL (English as a second language) teacher is to understand what my students struggle with in order to give necessary assistance where it is needed. This thesis has therefore unquestionably offered some valuable insight into this as I have gained knowledge of both second language acquisition in general and more specifically about different theories and hypotheses that explain both morphosyntactic problems and subject-verb agreement problems in different manners. By using this insight as a teacher, I think it will help me greatly with understanding why my students might struggle with different grammatical aspects.

Secondly, when writing this thesis I have looked at Norwegian and English Tense and Agreement from a comparative aspect. Seeing that there exists a universal consensus among linguists that L1 influences the acquisition of an L2, I think it is very important for an ESL teacher to know where both the differences and the similarities between the languages lie in order to fully understand why students make the errors they do. During this writing experience I have thus gained helpful knowledge about both Norwegian and English syntax which undoubtfully will help me in my upcoming teacher career.

Furthermore, I have always found it very important to be up to date on relevant research and literature in the pedagogical field, especially now when I am embarking on my profession as an ESL teacher. By writing this thesis, I have gained a lot of knowledge of both how to conduct a study as well as how to interpret scientific findings better. I firmly believe this will help me as a teacher as it makes research in general more accessible. By reading several research papers, I have also seen the value in varying one`s methods in research, which I also believe is necessary in teaching.

Overall, by working with this thesis I have gained a clearer view of what should be highlighted when teaching subject-verb agreement, and how this should be done in a way that will make it more manageable for the L1 learner. The thesis accentuates the value of exposure to the target language, which supports the predictions I had before writing this thesis.

Appendix II: Consent form

Would you like to participate in a research project on sentence processing in L2 English?

Background and purpose
This project is part of an MA thesis carried out by Mali Kokvoll at the Norwegian University of Science and Technology (NTNU). The purpose of the study is to look at the processing of English sentences by native Norwegian speakers and assess intuitions Norwegians have about English.

Procedure
If you participate, you will answer some basic questions about your language background and age, and then you will judge English sentences. The sentences will be presented one at a time and word-by-word. After reading each sentence, you will be asked to answer two questions about it. The first question will ask you to evaluate whether or not you think the English sentence was acceptable, and the second will how confident you are in your judgment.

Voluntary participation
Participation in this project is completely voluntary. You can freely choose to stop participating without giving an explanation by simply quitting the task before completion. There will be no negative consequences should you choose not to participate or to withdraw your consent.

Your information and data
You will not be asked to provide any personally identifiable information, so participation is completely anonymous. No potentially identifiable information will be published in the thesis.

The program used to run the study, IbexFarm, only collects data for participants who have completed the study. In order to process your results, your IP address will be stored temporarily by the software. When you complete the survey, the software anonymizes your data so it can't be traced back to your IP address. Your data will be stored on a password-protected server hosted by NTNU and will only be handled by me and the supervisors of the project. The project is scheduled to finish May 15, 2021, at which point the data will be removed from the server. Data will only be used for research purposes.

What gives us the right to process your personal data?
We will process your personal data based on your consent based on agreements with NTNU, NSD – The Norwegian Centre for Research Data AS has determined that the processing of personal data in this project is in accordance with data protection legislation.

Contact information
If you have any questions about the project or the information provided in this form, or if you want to exercise your rights, please contact me by e-mail: maliko@stud.ntnu.no. You may also contact Associate Professor Dave Kush (supervisor) at dave.kush@ntnu.no. If you have any questions regarding NSD and their evaluation of this project, you may contact them at personverntjenester@nsd.no or by phone: 55 58 21 17

Consent

I have received and understood the information about the project. By checking the box below, I affirm that:

I consent to participate in the project described on this page.

Appendix III: Background information form

Please complete this short questionnaire about your linguistic background before beginning the experiment. Please answer the questions accurately.

Age	[ ]
Gender:	○ Female ○ Male ○ Other ○ Prefer not to answer
Is Norwegian your native language?:	○ Yes ○ No
Did you grow up speaking English in your household?	○ Yes ○ No
When did you start learning English?	○ Birth - 4 years old ○ 4 - 6 years old ○ 6 - 10 years old ○ 10 - 15 years old ○ After age 15
Did you start learning English at school?	○ Yes ○ No
Do you regularly speak English with family members?	○ Yes ○ No
Did you grow up speaking a language other than Norwegian/Swedish/Danish in your household?	○ Yes ○ No
Do you speak a language other than Norwegian/Swedish/Danish fluently?	○ Yes ○ No

Are you currently studying at University?	○ Yes ○ No
What is/was your area of study at University?	○ Not Applicable (don't/didn't study at Uni) ○ Norwegian ○ English ○ Another Foreign Language ○ Another Social Science ○ Natural Science or Mathematics ○ Other
Approximately how many hours a week do you speak English, on average?	○ 0-1 ○ 2-5 ○ 5-10 ○ 10-15 ○ 15+
Approximately how many hours a week do you spend consuming English media, on average? (e.g. text, TV/videos, podcasts, lectures, music)	○ 0-1 ○ 2-5 ○ 5-10 ○ 10-15 ○ 15+

→ Click here to continue

Appendix IV: Items

	Factor 1	Factor 2	Factor 3	Subject	Relative Clause	Test word	Post-test
1	Sg	aux	G	The kid	that should listen	is	reading.
1	Sg	aux	UnG	The kid	that should listen	are	reading.
1	Pl	aux	UnG	The kids	that should listen	is	reading.
1	Pl	aux	G	The kids	that should listen	are	reading.
1	Sg	main	G	The kid	that should listen	reads	fast.
1	Sg	main	UnG	The kid	that should listen	read	fast.
1	Pl	main	G	The kids	that should listen	read	fast.
1	Pl	main	Ung	The kids	that should listen	reads	fast.
2	Sg	aux	G	The man	who enjoyed sunlight	is	walking.
2	Sg	aux	UnG	The man	who enjoyed sunlight	are	walking.
2	Pl	aux	UnG	The men	who enjoyed sunlight	is	walking.
2	Pl	aux	G	The men	who enjoyed sunlight	are	walking.
2	Sg	main	G	The man	who enjoyed sunlight	walks	a lot.
2	Sg	main	UnG	The man	who enjoyed sunlight	walk	a lot.
2	Pl	main	G	The men	who enjoyed sunlight	walk	a lot.
2	Pl	main	UnG	The men	who enjoyed sunlight	walks	a lot.
3	Sg	aux	G	The guest	that arrived yesterday	is	complaining.
3	Sg	aux	UnG	The guest	that arrived yesterday	are	complaining.
3	Pl	aux	UnG	The guests	that arrived yesterday	is	complaining.
3	Pl	aux	G	The guests	that arrived yesterday	are	complaining.
3	Sg	main	G	The guest	that arrived yesterday	complains	a lot.
3	Sg	main	UnG	The guest	that arrived yesterday	complain	a lot.
3	Pl	main	UnG	The guests	that arrived yesterday	complains	a lot.
3	Pl	main	G	The guests	that arrived yesterday	complain	a lot.
4	Sg	aux	G	The girl	that should save	is	spending money.
4	Sg	aux	UnG	The girl	that should save	are	spending money.
4	Pl	aux	UnG	The girls	that should save	is	spending money.
4	Pl	aux	G	The girls	that should save	are	spending money.

4	Sg	main	G	The girl	that should save	spends	lots of money.
4	Sg	main	UnG	The girl	that should save	spend	lots of money.
4	Pl	main	UnG	The girls	that should save	spends	lots of money.
4	Pl	main	G	The girls	that should save	spend	lots of money.
5	Sg	aux	G	The kid	who loved Christmas	is	wishing for snow.
5	Sg	aux	UnG	The kid	who loved Christmas	are	wishing for snow.
5	Pl	aux	Ung	The kids	who loved Christmas	is	wishing for snow.
5	Pl	aux	G	The kids	who loved Christmas	are	wishing for snow.
5	Sg	main	G	The kid	who loved Christmas	wishes	for snow.
5	Sg	main	UnG	The kid	who loved Christmas	wish	for snow.
5	Pl	main	UnG	The kids	who loved Christmas	wishes	for snow.
5	Pl	main	G	The kids	who loved Christmas	wish	for snow.
6	Sg	aux	G	The woman	that should rest	is	running.
6	Sg	aux	UnG	The woman	that should rest	are	running.
6	Pl	aux	UnG	The women	that should rest	is	running.
6	Pl	aux	G	The women	that should rest	are	running.
6	Sg	main	G	The woman	that should rest	runs	fast.
6	Sg	main	UnG	The woman	that should rest	run	fast.
6	Pl	main	UnG	The women	that should rest	runs	fast.
6	Pl	main	G	The women	that should rest	run	fast.
7	Sg	aux	G	The boy	that had fallen	is	yelling.
7	Sg	aux	UnG	The boy	that had fallen	are	yelling.
7	Pl	aux	UnG	The boys	that had fallen	is	yelling.
7	Pl	aux	G	The boys	that had fallen	are	yelling.
7	Sg	main	G	The boy	that had fallen	yells	loudly.
7	Sg	main	UnG	The boy	that had fallen	yell	loudly.
7	Pl	main	UnG	The boys	that had fallen	yells	loudly.
7	Pl	main	G	The boys	that had fallen	yell	loudly.
8	Sg	aux	G	The painter	who lived nearby	is	working.
8	Sg	aux	UnG	The painter	who lived nearby	are	working.
8	Pl	aux	UnG	The painters	who lived nearby	is	working.
8	Pl	aux	G	The painters	who lived nearby	are	working.
8	Sg	main	G	The painter	who lived nearby	works	hard.
8	Sg	main	UnG	The painter	who lived nearby	work	hard.

8	Pl	main	UnG	the painters	who lived nearby	works	hard.
8	Pl	main	G	the painters	who lived nearby	work	hard.
9	Sg	aux	G	The horse	that had competed	is	eating hay.
9	Sg	aux	UnG	The horse	that had competed	are	eating hay.
9	Pl	aux	UnG	The horses	that had competed	is	eating hay.
9	Pl	aux	G	The horses	that had competed	are	eating hay.
9	Sg	main	G	The horse	that had competed	eats	lots of hay.
9	Sg	main	UnG	The horse	that had competed	eat	lots of hay.
9	Pl	main	UnG	The horses	that had competed	eats	lots of hay.
9	Pl	main	G	The horses	that had competed	eat	lots of hay.
10	Sg	aux	G	The student	that had failed	is	succeeding in classes now.
10	Sg	aux	UnG	The student	that had failed	are	succeeding in classes now.
10	Pl	aux	UnG	The students	that had failed	is	succeeding in classes now.
10	Pl	aux	G	The students	that had failed	are	succeeding in classes now.
10	Sg	main	G	The student	that had failed	succeeds	in classes now.
10	Sg	main	UnG	The student	that had failed	succeed	in classes now.
10	Pl	main	UnG	The students	that had failed	succeeds	in classes now.
10	Pl	main	G	The students	that had failed	succeed	in classes now.
11	Sg	aux	G	The pianist	that should practice	is	playing everyday.
11	Sg	aux	UnG	The pianist	that should practice	are	playing everyday
11	Pl	aux	UnG	The pianists	that should practice	is	playing everyday
11	Pl	aux	G	The pianists	that should practice	are	playing everyday
11	Sg	main	G	The pianist	that should practice	plays	everyday
11	Sg	main	UnG	The pianist	that should practice	play	everyday
11	Pl	main	UnG	The pianists	that should practice	plays	everyday
11	Pl	main	G	The pianists	that should practice	play	everyday
12	Sg	aux	G	The cat	that had attacked	is	playing outside the house.
12	Sg	aux	UnG	The cat	that had attacked	are	playing outside the house.
12	Pl	aux	UnG	The cats	that had attacked	is	playing outside the house.
12	Pl	aux	G	The cats	that had attacked	are	playing outside the house.
12	Sg	main	G	The cat	that had attacked	plays	outside the house.

54

12	Sg	main	UnG	The cat	that had attacked	play	outside the house.
12	Pl	main	UnG	The cats	that had attacked	plays	outside the house.
12	Pl	main	G	The cats	that had attacked	play	outside the house.
13	Sg	aux	G	The policeman	that loved working	is	planning to retire.
13	Sg	aux	UnG	The policeman	that loved working	are	planning to retire.
13	Pl	aux	UnG	The policemen	that loved working	is	planning to retire.
13	Pl	aux	G	The policemen	that loved working	are	planning to retire.
13	Sg	main	G	The policeman	that loved working	plans	to retire.
13	Sg	main	UnG	The policeman	that loved working	plan	to retire.
13	Pl	main	UnG	The policemen	that loved working	plans	to retire.
13	Pl	main	G	The policemen	that loved working	plan	to retire.
14	Sg	aux	G	The bride	who hated shopping	is	complaining about everything.
14	Sg	aux	UnG	The bride	who hated shopping	are	complaining about everything.
14	Pl	aux	UnG	The brides	who hated shopping	is	complaining about everything.
14	Pl	aux	G	The brides	who hated shopping	are	complaining about everything.
14	Sg	main	G	The bride	who hated shopping	complains	about everything.
14	Sg	main	UnG	The bride	who hated shopping	complain	about everything.
14	Pl	main	UnG	The brides	who hated shopping	complains	about everything.
14	Pl	main	G	The brides	who hated shopping	complain	about everything.
15	Sg	aux	G	The mother	who delivered yesterday	is	nursing her baby
15	Sg	aux	UnG	The mother	who delivered yesterday	are	nursing her baby
15	Pl	aux	UnG	The mothers	who delivered yesterday	is	nursing her baby
15	Pl	aux	G	The mothers	who delivered yesterday	are	nursing her baby
15	Sg	main	G	The mother	who delivered yesterday	nurses	her baby
15	Sg	main	UnG	The mother	who delivered yesterday	nurse	her baby
15	Pl	main	UnG	The mothers	who delivered yesterday	nurses	her baby
15	Pl	main	G	The mothers	who delivered yesterday	nurse	her baby
16	Sg	aux	G	The nurse	that should stay	is	leaving today.
16	Sg	aux	UnG	The nurse	that should stay	are	leaving today.
16	Pl	aux	UnG	The nurses	that should stay	is	leaving today.
16	Pl	aux	G	The nurses	that should stay	are	leaving today.

16	Sg	main	G	The nurse	that should stay	leaves	today.
16	Sg	main	UnG	The nurse	that should stay	leave	today.
16	Pl	main	UnG	The nurses	that should stay	leaves	today.
16	Pl	main	G	The nurses	that should stay	leve	today.
17	Sg	aux	G	The lion	that hurt itself	is	roaring at the visitors.
17	Sg	aux	UnG	The lion	that hurt itself	are	roaring at the visitors.
17	Pl	aux	UnG	The lions	that hurt itself	is	roaring at the visitors.
17	Pl	aux	G	The lions	that hurt itself	are	roaring at the visitors.
17	Sg	main	G	The lion	that hurt itself	roars	at every visitor.
17	Sg	main	UnG	The lion	that hurt itself	roar	at every visitor.
17	Pl	main	UnG	The lions	that hurt itself	roars	at every visitor.
17	Pl	main	G	The lions	that hurt itself	roar	at every visitor.
18	Sg	aux	G	The teacher	that had resigned	is	working again.
18	Sg	aux	UnG	The teacher	that had resigned	are	working again.
18	Pl	aux	UnG	The teachers	that had resigned	is	working again.
18	Pl	aux	G	The teachers	that had resigned	are	working again.
18	Sg	main	G	The teacher	that had resigned	works	again.
18	Sg	main	UnG	The teacher	that had resigned	work	again.
18	Pl	main	UnG	The teachers	that had resigned	works	again.
18	Pl	main	G	The teachers	that had resigned	work	again.
19	Sg	aux	G	The trainer	that might quit	is	training lazy clients.
19	Sg	aux	UnG	The trainer	that might quit	are	training lazy clients.
19	Pl	aux	UnG	The trainers	that might quit	is	training lazy clients.
19	Pl	aux	G	The trainers	that might quit	are	training lazy clients.
19	Sg	main	G	The trainer	that might quit	trains	lots of lazy clients.
19	Sg	main	UnG	The trainer	that might quit	train	lots of lazy clients.
19	Pl	main	UnG	The trainers	that might quit	trains	lots of lazy clients.
19	Pl	main	G	The trainers	that might quit	train	lots of lazy clients.
20	Sg	aux	G	The dog	that slept alot	is	participating in many contests.
20	Sg	aux	UnG	The dog	that slept alot	are	participating in many contests.
20	Pl	aux	UnG	The dogs	that slept alot	is	participating in many contests.
20	Pl	aux	G	The dogs	that slept alot	are	participating in many contests.
20	Sg	main	G	The dog	that slept alot	participates	in many contests.
20	Sg	main	UnG	The dog	that slept alot	participate	in many contests.
20	Pl	main	UnG	The dogs	that slept alot	participates	in many contests.

20	Pl	main	G	The dogs	that slept alot	participate	in many contests.
21	Sg	aux	G	The king	who should rule	is	avoiding his responsibilities.
21	Sg	aux	UnG	The king	who should rule	are	avoiding his responsibilities.
21	Pl	aux	UnG	The kings	who should rule	is	avoiding his responsibilities.
21	Pl	aux	G	The kings	who should rule	are	avoiding his responsibilities.
21	Sg	main	G	The king	who should rule	avoids	his responsibilities.
21	Sg	main	UnG	The king	who should rule	avoid	his responsibilities.
21	Pl	main	UnG	The kings	who should rule	avoids	his responsibilities.
21	Pl	main	G	The kings	who should rule	avoid	his responsibilities.
22	Sg	aux	G	The panda	which might fall	is	sleeping near the ledge.
22	Sg	aux	UnG	The panda	which might fall	are	sleeping near the ledge.
22	Pl	aux	UnG	The pandas	which might fall	is	sleeping near the ledge.
22	Pl	aux	G	The pandas	which might fall	are	sleeping near the ledge.
22	Sg	main	G	The panda	which might fall	sleeps	near the ledge.
22	Sg	main	UnG	The panda	which might fall	sleep	near the ledge.
22	Pl	main	UnG	The pandas	which might fall	sleeps	near the ledge.
22	Pl	main	G	The pandas	which might fall	sleep	near the ledge.
23	Sg	aux	G	The commentator	that loved soccer	is	supporting Manchester United.
23	Sg	aux	UnG	The commentator	that loved soccer	are	supporting Manchester United.
23	Pl	aux	UnG	The commentators	that loved soccer	is	supporting Manchester United.
23	Pl	aux	G	The commentators	that loved soccer	are	supporting Manchester United.
23	Sg	main	G	The commentator	that loved soccer	supports	Manchester United.
23	Sg	main	UnG	The commentator	that loved soccer	support	Manchester United.
23	Pl	main	UnG	The commentators	that loved soccer	supports	Manchester United.
23	Pl	main	G	The commentators	that loved soccer	support	Manchester United.
24	Sg	aux	G	The child	that had surgery	is	starting school tomorrow.
24	Sg	aux	UnG	The child	that had surgery	are	starting school tomorrow.
24	Pl	aux	UnG	The children	that had surgery	is	starting school tomorrow.
24	Pl	aux	G	The children	that had surgery	ate	starting school tomorrow.
24	Sg	main	G	The child	that had surgery	starts	school tomorrow.

24	Sg	main	UnG	The child	that had surgery	start	school tomorrow.
24	Pl	main	UnG	The children	that had surgery	starts	school tomorrow.
24	Pl	main	G	The children	that had surgery	start	school tomorrow.
25	Sg	aux	G	The father	who played hockey	is	watching the match.
25	Sg	aux	UnG	The father	who played hockey	are	watching the match.
25	Pl	aux	UnG	The fathers	who played hockey	is	watching the match.
25	Pl	aux	G	The fathers	who played hockey	are	watching the match.
25	Sg	main	G	The father	who played hockey	watches	all the matches.
25	Sg	main	UnG	The father	who played hockey	watch	all the matches.
25	Pl	main	UnG	The fathers	who played hockey	watches	all the matches.
25	Pl	main	G	The fathers	who played hockey	watch	all the matches.
26	Sg	aux	G	The toddler	who started walking	is	smiling at me.
26	Sg	aux	UnG	The toddler	who started walking	are	smiling at me.
26	Pl	aux	UnG	The toddlers	who started walking	is	smiling at me.
26	Pl	aux	G	The toddlers	who started walking	are	smiling at me.
26	Sg	main	G	The toddler	who started walking	smiles	all the time.
26	Sg	main	UnG	The toddler	who started walking	smile	all the time.
26	Pl	main	UnG	The toddlers	who started walking	smiles	all the time.
26	Pl	main	G	The toddlers	who started walking	smile	all the time.
27	Sg	aux	G	The norwegian	who knew English	is	talking quickly.
27	Sg	aux	UnG	The norwegian	who knew English	are	talking quickly.
27	Pl	aux	UnG	The norwegians	who knew English	is	talking quickly.
27	Pl	aux	G	The norwegians	who knew English	are	talking quickly.
27	Sg	main	G	The norwegian	who knew English	talks	quickly.
27	Sg	main	UnG	The norwegian	who knew English	talk	quickly.
27	Pl	main	UnG	The norwegians	who knew English	talks	quickly.
27	Pl	main	G	The norwegians	who knew English	talk	quickly.
28	Sg	aux	G	The neighbor	that enjoyed carpentry	is	building fences.
28	Sg	aux	UnG	The neighbor	that enjoyed carpentry	are	building fences.

28	Pl	aux	UnG	The neighbors	that enjoyed carpentry	is	building fences.
28	Pl	aux	G	The neighbors	that enjoyed carpentry	are	building fences.
28	Sg	main	G	The neighbor	that enjoyed carpentry	builds	many fences.
28	Sg	main	UnG	The neighbor	that enjoyed carpentry	build	many fences.
28	Pl	main	UnG	The neighbors	that enjoyed carpentry	builds	many fences.
28	Pl	main	G	The neighbors	that enjoyed carpentry	build	many fences.
29	Sg	aux	G	The clown	that looked funny	is	scaring everyone.
29	Sg	aux	UnG	The clown	that looked funny	are	scaring everyone.
29	Pl	aux	UnG	The clowns	that looked funny	is	scaring everyone.
29	Pl	aux	G	The clowns	that looked funny	are	scaring everyone.
29	Sg	main	G	The clown	that looked funny	scares	everyone.
29	Sg	main	UnG	The clown	that looked funny	scare	everyone.
29	Pl	main	UnG	The clowns	that looked funny	scares	everyone.
29	Pl	main	G	The clowns	that looked funny	scare	everyone.
30	Sg	aux	G	The nanny	that often worked	is	helping the child.
30	Sg	aux	UnG	The nanny	that often worked	are	helping the child.
30	Pl	aux	UnG	The nannies	that often worked	is	helping the child.
30	Pl	aux	G	The nannies	that often worked	are	helping the child.
30	Sg	main	G	The nanny	that often worked	helps	the child.
30	Sg	main	UnG	The nanny	that often worked	help	the child.
30	Pl	main	UnG	The nannies	that often worked	help	the child.
30	Pl	main	G	The nannies	that often worked	helps	the child.
31	Sg	aux	G	The athlete	who never rested	is	losing the race.
31	Sg	aux	UnG	The athlete	who never rested	are	losing the race.
31	Pl	aux	UnG	The athletes	who never rested	is	losing the race.
31	Pl	aux	G	The athletes	who never rested	are	losing the race.
31	Sg	main	G	The athlete	who never rested	loses	many races.
31	Sg	main	UnG	The athlete	who never rested	lose	many races.
31	Pl	main	UnG	The athletes	who never rested	loses	many races.

31	Pl	main	G	The athletes	who never rested	lose	many races.
32	Sg	aux	G	The musician	that had performed	is	missing her family.
32	Sg	aux	UnG	The musician	that had performed	are	missing her family.
32	Pl	aux	UnG	The musicians	that had performed	is	missing her family.
32	Pl	aux	G	The musicians	that had performed	are	missing her family.
32	Sg	main	G	The musician	that had performed	misses	her family.
32	Sg	main	UnG	The musician	that had performed	miss	her family.
32	Pl	main	UnG	The musicians	that had performed	misses	her family.
32	Pl	main	G	The musicians	that had performed	miss	her family.

Appendix V: Fillers

Items	Filler Type	Bad/Good	NP	RC	VP	Test-phrase
Item variation 1						
1	Many/Much	Bad	The grizzly	that was awake	chewed	many food
1	Many/Much	Good	The grizzly	that was awake	chewed	many berries
2	Many/Much	Bad	The grandmother	who always smiled	drank	much bananas
2	Many/Much	Good	The grandmother	who always smiled	drank	much water
3	Many/Much	Bad	The teacher	that worked a lot	helped	much children.
3	Many/Much	Good	The teacher	that worked a lot	helped	many children.
4	Many/Much	Bad	The pilot	that enjoyed heights	flies	many plane.
4	Many/Much	Good	The pilot	that enjoyed heights	flies	many planes.
5	Many/Much	Bad	The chef	that loves food	tastes	many glasses daily.
5	Many/Much	Good	The chef	that loves food	tastes	many dishes daily.
6	Many/Much	Bad	The professor	who is old	reads	much books.
6	Many/Much	Good	The professor	who is old	reads	many books.
7	Many/Much	Bad	The students	that structures well	have	much clock
7	Many/Much	Good	The students	that structures well	have	much free time
8	Many/Much	Bad	The children	that had played	ate	much food
8	Many/Much	Good	The children	that had played	ate	much apples

Item variation 2						
9	Selection Violation	Bad	The dog	that was sick	needed	a lot of pills.
9	Selection Violation	Good	The dog	that was sick	needed	a lot of man.
10	Selection Violation	Bad	The millionaire	that earned a lot	bought	a cash.
10	Selection Violation	Good	The millionaire	that earned a lot	bought	a house.
11	Selection Violation	Bad	The child	who lived nearby	listened	to hearing.
11	Selection Violation	Good	The child	who lived nearby	listened	to music.
12	Selection Violation	Bad	The pupil	that teachers love	sharpened	all the books.
12	Selection Violation	Good	The pupil	that teachers love	sharpened	all the pencils.
13	Selection Violation	Bad	The tiger	that is wild	attacks	many thoughts.
13	Selection Violation	Good	The tiger	that is wild	attacks	many antelopes.
14	Selection Violation	Bad	The student	that needed to study	read	a scissor.
14	Selection Violation	Good	The student	that needed to study	read	a book.

15	Selection Violation	Bad	The boy	who was hungry	ate	pizza.
15	Selection Violation	Good	The boy	who was hungry	ate	clouds.
16	Selection Violation	Bad	The father	that enjoyed speed	drove	a car.
16	Selection Violation	Good	The father	that enjoyed speed	drove	a tire.

Item variation 3						
17	Count/mass	Bad	A shopper	found a wallet	that contained	three coin.
17	Count/mass	Good	A shopper	found a wallet	that contained	three bills.
18	Count/mass	Bad	The cook	prepared a recipe	that called for	lots of milks.
18	Count/mass	Good	The cook	prepared a recipe	that called for	lots of milk.
19	Count/mass	Bad	The zookeeper	cleaned the tank	that was home to	five snake.
19	Count/mass	Good	The zookeeper	cleaned the tank	that was home to	five fish.
20	Count/mass	Bad	The lifeguard	walked on the path	that was covered	with sands.
20	Count/mass	Good	The lifeguard	walked on the path	that was covered	with sand.
21	Count/mass	Bad	The therapist	helped the patient	that asked for	few advice.
21	Count/mass	Good	The therapist	helped the patient	that asked for	little advice.
22	Count/mass	Bad	The programmer	broke the computer	that was compatible with	few software.
22	Count/mass	Good	The programmer	broke the computer	that was compatible with	some software.
23	Count/mass	Bad	The model	tried on a dress	that was made by	a designer
23	Count/mass	Good	The model	tried on a dress	that was made by	a designers
24	Count/mass	Bad	The soccer player	trained for the game	that was scheduled	in one day
24	Count/mass	Good	The soccer player	trained for the game	that was scheduled	in one days

Item variation 4						**reflexives (Test-phrase)**	
25	Reflexives	Bad	The students	that	blamed	himself	failed the exam.
25	Reflexives	Good	The students	that	blamed	themselves	failed the exam.
26	Reflexives	Bad	The singer	that	prepared	itself	won the competition.
26	Reflexives	Good	The singer	that	prepared	herself	won the competition.
27	Reflexives	Bad	The girl	that	enjoys	himself	got a new boyfriend.
27	Reflexives	Good	The girl	that	enjoys	herself	got a new boyfriend.
28	Reflexives	Bad	The man	that	hurt	herself	called to the women.
28	Reflexives	Good	The man	that	hurt	himself	called to the women.
29	Reflexives	Bad	The nurse	that	introduced	themselves	walked into the room.

#	Category						
29	Reflexives	Good	The nurse	that	introduced	herself	walked into the room.
30	Reflexives	Bad	The artists	that	painted	itself	were known all over.
30	Reflexives	Good	The artists	that	painted	themselves	were known all over.
31	Reflexives	Bad	The king	that	loved	itself	was hated by the people.
31	Reflexives	Good	The king	that	loved	himself	was hated by the people.
32	Reflexives	Bad	The cat	that	licked	themselves	snuggled the girl.
32	Reflexives	Good	The cat	that	cut	itself	snuggled the girl.

Item Variation 5					
17	Fronted object	Good	The man	that she saw	left.
17	Fronted object	Bad	The man	that she saw him	left.
18	Fronted object	Good	The dog	that he owned	is friendly.
18	Fronted object	Bad	The dog	that he owned it	is friendly.
19	Fronted object	Good	The kid	that they nursed	cried.
19	Fronted object	Bad	The kid	that they nursed him	cried.
20	Fronted object	Good	The lady	that he loved	ran away.
20	Fronted object	Bad	The lady	that he loved lady	ran away.
21	Fronted object	Good	The musicians	that the crowd loved	played well.
21	Fronted object	Bad	The musicians	that the crowd loved them	played well.
22	Fronted object	Good	The boy	that she knew	smiled.
22	Fronted object	Bad	The boy	that she knew him	smiled.
23	Fronted object	Good	The singer	that they listened to	died.
23	Fronted object	Bad	The singer	that they listened to him	died.
24	Fronted object	Good	The mother	that he knew	came early.
24	Fronted object	Bad	The mother	that he knew her	came early.

Item Variation 6								
1	Filler-gap resolution	Bad	The inspector	wondered	what	the thief	had taken	the necklace	that night.
1	Filler-gap resolution	Good	The inspector	wondered	what	the thief	had taken		that night.
2	Filler-gap resolution	Bad	The translator	asked	who	the ambassador	avoided	the diplomat	at the gala.
2	Filler-gap resolution	Good	The translator	asked	who	the ambassador	avoided		at the gala.
3	Filler-gap resolution	Bad	The baker	wondered	what	the customer	had bought	the bread	that day.
3	Filler-gap resolution	Good	The baker	wondered	what	the customer	had bought		that day.
4	Filler-gap resolution	Bad	The clown	asked	what	the child	wanted	the balloon	at the party.

4	Filler-gap resolution	Good	The clown	asked	what	the child	wanted		at the party.
5	Filler-gap resolution	Bad	The receptionist	wondered	who	the guest	looked for	his wife	yesterday.
5	Filler-gap resolution	Good	The receptionist	wondered	who	the guest	looked for		yesterday.
6	Filler-gap resolution	Bad	The pilot	asked	what	the stewardess	needed	a life west	for the presentation.
6	Filler-gap resolution	Good	The pilot	asked	what	the stewardess	needed		for the presentation.
7	Filler-gap resolution	Bad	The lawyer	wondered	who	the judge	sentenced	the accused offender	that day.
7	Filler-gap resolution	Good	The lawyer	wondered	who	the judge	sentenced		that day.
8	Filler-gap resolution	Bad	The veterinarian	asked	what	the dog owner	fed the dog	meatloaf	for dinner.
8	Filler-gap resolution	Good	The veterinarian	asked	what	the dog owner	fed the dog		for dinner.
9	Filler-gap resolution	Bad	The teacher	wondered	where	the student	put the laptop	in the drawer	yesterday.
9	Filler-gap resolution	Good	The teacher	wondered	where	the student	put the laptop		yesterday.
10	Filler-gap resolution	Bad	The swimmer	wondered	who	the lifeguard	saved	a little child	at the swimming pool.
10	Filler-gap resolution	Good	The swimmer	wondered	who	the lifeguard	saved		at the swimming pool.
11	Filler-gap resolution	Bad	The athlete	asked	what	the coach	had planned	intervals	for today´s session.
11	Filler-gap resolution	Good	The athlete	asked	what	the coach	had planned		for today´s session.
12	Filler-gap resolution	Bad	The firefighter	wondered	what	the house owners	had lost	everything	in the fire.
12	Filler-gap resolution	Good	The firefighter	wondered	what	the house owners	had lost		in the fire.