UiT The Arctic University of Norway

Faculty of Humanities, Social Sciences and Education

Teaching agreement through grammar instruction in the lower secondary school

A comparison between implicit and explicit grammar instruction in L1 Norwegian learners of L2 English

Eirik Skaug

Master thesis in English language and education, ENG-3982, May 2022

Table of Contents

List of tables

List of figures

Acknowledgements

First, I would like to express my most heartfelt gratitude to my supervisors, Natalia Mitrofanova and Christopher Loe Olsen, for providing essential insights and guidance during my writing process.

I would also like to thank my family for supporting me and always being there when I needed motivation and advice.

Lastly, my sincerest thanks to the people I have gotten to know through my five years of study at the Lektor program in Tromsø. The last few months would not have been as tolerable without you.

1 Introduction

This thesis investigates L1 Norwegian's knowledge of L2 English subject-verb agreement in the lower secondary school. More specifically, the groups examined in the present study are tested using acceptability judgement tasks to assess their accuracy in judging correct and incorrect sentences containing subject-verb agreement. Furthermore, the study will explore the effects of grammar instruction through two different interventions, where the effects of the interventions are measured through two post-tests. The post-tests are conducted after the intervention, the first is used to measure the short-term effects and the second, the delayed post-test, is used to measure long-term effects.

Unlike Norwegian, English agreement uses the third person singular "*-s*" as an agreement marker in subject-verb agreement. Norwegian has no overt agreement marker, a contrast which might be part of the reason as to why L1 Norwegians find this grammatical construction challenging. Consequently, the problem might be related to crosslinguistic influence between the two languages, with L1 Norwegian causing negative transfer.

The third person singular "*-s*" construction is a part of English functional morphology, which makes up a language's formal grammatical features. Functional morphology is hypothesized as being a locus of difficulty for language learners, as this is where much of the differences between languages are located (Jensen et al., 2020, 4). Theories like the Bottleneck Hypothesis (Slabakova, 2008) and Syntax-Before-Morphology (White, 2003) maintain this view, considering morphology a challenging part of language acquisition. Studies investigating Norwegian acquisition of English subject-verb agreement include Jensen et al. (2019), which examined the bottleneck hypothesis, and Garshol (2018). These studies will be discussed throughout this thesis.

The present study compares two teaching methods to measure their effects on subject-verb agreement acquisition in instructional settings. The teaching methods, Presentation Practice Production (PPP) and Task Based Language Teaching (TBLT) use explicit and implicit instruction, respectively.

Limited research has been conducted on the effects of instruction on English subject-verb agreement, so this study can shed light on possible effects this might have on acquisition. Furthermore, no study has specifically compared the effects of these teaching methods (PPP and TBLT) on L1 Norwegian L2 English subject-verb agreement acquisition. The study

emphasizes the pedagogical dimension of language acquisition by including the instructional intervention; thus, the findings are also discussed in a classroom context.

The study's design was motivated by a perceived lack of applied second language research in Norwegian classrooms. There seems to be a divide between instructional practitioners and researchers that makes communication and application more complicated than necessary (Toth & Moranski, 2018, 74). Research should be available for practitioners to provide the most effective language instruction. The problem is likely a two-way street, and both academia, instructors and educational institutions are to blame for the limited communication. Nevertheless, studies like the present one, where both linguistic phenomena and instructional methods are explored, might make L2 research more available for instruction.

This thesis will seek to answer four research questions:

> **RQ 1:** Do Norwegian L1 learners of English have little knowledge of the third person singular "-s"?
>
> **RQ 2:** Is there any variation in accuracy on the sub-categories; local singular agreement, local plural agreement, long-distance singular agreement and long-distance plural agreement among L1 Norwegian L2 English learners?
>
> **RQ 3:** Does instruction affect learner's knowledge of subject verb agreement?
>
> **RQ 4:** Do different types of instruction affect learner's knowledge of subject-verb agreement differently?

The study employed three acceptability judgment tasks to answer these questions: A pre-test before the pedagogical intervention, a post-test one day after the intervention, and a delayed post-test two months after the intervention. The intervention was different for the two groups, one using explicit grammar teaching (PPP) and one using implicit grammar teaching (TBLT). Each intervention lasted 60 minutes. Additionally, the study used a subset of the Oxford Proficiency test to assess the participant's English proficiency and a background questionnaire.

The acceptability judgement tasks were also used to measure the student's accuracy on specific sub-conditions of subject-verb agreement:

1. Local singular agreement

2. Local plural agreement
3. Long-distance singular agreement
4. Long-distance plural agreement

The thesis is divided into the following sections: 2 describes the background, which in turn is divided into three parts; Empirical background (2.1), theoretical framework: GenSLA (2.2) and theoretical framework: ISLA (2.3). 3 further describes the research questions and predictions, 4 describes the methodology, 5 display the results, 6 presents some of the limitations of the study, 7 discusses the findings and tries to answer the research questions and, finally, 8 provides a conclusion to the thesis.

2 Background

This section, describing the background used in this thesis, is split into three sub-sections. The first part (2.1) showcases empirical studies on subject-verb agreement acquisition in L2 English learners and explains this linguistic phenomenon in detail. The following two sub-sections (2.2 and 2.3) explore the two theoretical frameworks that are used; generative second language acquisition (GenSLA) and instructional second language acquisition (ISLA). The two are distinct, mainly because the former is a purely theoretical linguistic field, while the latter discusses applied linguistic and didactic theory.

2.1 Empirical background

Research on L2 acquisition has shown that some properties of languages are more complicated to acquire than others (Jensen et al., 2019, 4). For L1 Norwegian learners of English, one of these properties is subject-verb agreement, especially the use of the third person singular "*-s*". Firstly, this section will provide data on L2 English agreement marking in general, and then, more specifically examine agreement marking in L1 Norwegian learners of L2 English. Following this, the differences between agreement marking in Norwegian and English are highlighted in 2.1.3.

In the following study, the group under investigation is adolescents (15-16 years); however, most of the data that has been collected on subject-verb agreement marking is from adult learners (Garshol, 2019, 40). Therefore, some of the studies that have been chosen for comparison are mainly from adult learners. On the other hand, these can be considered viable for contrast because of the typical English proficiency for Norwegian 15-16-year-olds. The reasoning behind this is that obligatory English instruction ends in the first year of upper secondary school in the Norwegian school system, and the participants of this study were halfway through their last year of lower secondary school when the tests were conducted. In this respect, they can be viewed as almost finished with their formal English education and, therefore close to the proficiency level of a typical Norwegian adult. Their proficiency is also reflected in test scores, which are discussed more closely in section 4. Furthermore, proficiency can vary significantly in the same age group; therefore, using age as the sole measurement of proficiency may be inaccurate.

According to theories on L2 acquisition, like The Bottleneck Hypothesis (Slabakova, 2008, 2016) and Syntax Before Morphology (Lardiere, 1998a; Lardiere, 1998b; White, 2003), which will be discussed later in this section, functional morphology is especially challenging to acquire for L2 learners. To further highlight this, the following studies which have examined the acquisition of functional morphology in L2 English learners are discussed.

2.1.1 Subject-verb agreement in adults

Jackson et al. (2018) focus on the effects of the language learner's L1 on production of subject-verb agreement in the L2. The study contrasted L1 Chinese L2 English learners and L1 Swedish L2 English learners, specifically comparing how L2 English proficiency and L1 morphosyntax affected English subject-verb agreement. The study problematizes that noun phrases (NP) often have ambiguous grammatical and conceptual functions, like the NP "scissors", which are conceptually singular, yet grammatically plural. According to this paper, this ambiguousness is a source of error in the L2 learner. In the study, the Chinese and Swedish participants had similar amounts of subject-verb agreement errors. The study also found that despite the participant's native-like performance in the English proficiency test, they still produced more subject-verb agreement errors than the English L1 baseline.

These results show that the same underlying mechanisms for acquiring subject-verb agreement are present, regardless of the learner's L1 base (Jackson, 2018, 917). Furthermore, the results show that subject-verb agreement is challenging to acquire relative to other linguistic features

2.1.2 Studies on acquisition of English agreement in Norwegian adolescents

In Jensen et al. (2019), 60 students aged between 11 and 18 years were tested using an acceptability judgement test. Like the present study, this study tested four sub-conditions: 1) Local agreement with a singular subject, 2) Local agreement with a plural subject, 3) Long-distance agreement with a plural subject and 4) Long-distance agreement with a plural subject, all in present subject-initial declarative clauses. As opposed to the acceptability judgement task used in the present study, this study measured the participant's accuracy using a Likert scale (1-4). However, 1 and 2 both represent grammatical judgement and 3 and 4 both represent ungrammatical judgement, making the scoring of the participant's judgement binary, like in this study. The study used a general English proficiency test as a reference point to see the relative difficulty of subject-verb agreement. Furthermore, the study also tested knowledge of V2-word

order to compare the difficulty between phenomena related to functional morphology and phenomena pertaining to core syntax.

The results in Jensen et al. (2019, 16-21) showed that accuracy was significantly higher in both sub-conditions testing for knowledge on V2-word order compared to all four sub-conditions testing for knowledge on subject-verb agreement. Among the four agreement sub-conditions, accuracy in local singular, long-distance singular and local plural were clustered together. At the same time, the participants had scored even lower on the sub-condition targeting long-distance plural agreement. The researchers confirm that the results support their hypothesis, stating that subject-verb agreement is a more persistent problem than V2-word movement in L2 acquisition of English. Based on the differences in scores on the four morphological sub-conditions, the researchers also conclude that plurality and additional lexical items between the subject and the verb (distance) cause difficulty in L2 learners. Furthermore, the study found that the participants tended to prefer sentences that contained the suffix third person singular "-s", a finding that the researchers account to overgeneralization.

Garshol (2018) performed a longitudinal case study using corpus data. The data was collected over a year, following two groups of upper-secondary students. One intervention group with 64 students was measured in the study, and one control group with 69 students. Their data was collected through texts the students admitted throughout the year during their English course; however, they accumulated at three intervention points (fall, midterm and spring).

The accumulated corpus only checked for general agreement errors and did not analyze specific sub-types. Nevertheless, this data is viable as it underscores the difficulty of acquiring English morphology by L1 Norwegians. The results showed no significant improvement in the intervention group over the year; however, the author prescribes this to the execution of the intervention and not to the design itself. Nevertheless, the results did show that the groups had higher error rates related to agreement than L2 English learners with other L1's (Garshol, 2018, 198). The data from both Jensen et al. (2019) and Garshol (2018) points toward difficulties acquiring agreement in L2 English on the part of L1 Norwegians, both specifically related to the third person singular "-s" (Jensen, 2019) and agreement in general (Garshol, 2018), showing that functional morphology might be a stumbling block for Norwegian learners. The following subsections will describe and compare the third-person singular "-s" construction in English to the Norwegian agreement construction and then present the theories mentioned above, the

bottleneck hypothesis (Slabakova, 2008) and syntax-before-morphology (White, 2003) that explain why acquiring this construction seems particularly challenging.

2.1.3 Subject-verb agreement

Subject-verb agreement in Norwegian and English differs concerning morphology. There are no overt morphemes marking for agreement between subject and verb in Norwegian, while English marks agreement overtly, however, only in the 3rd person singular, represented by adding the suffix ”-s”. These differences are further highlighted in the two following sub-sections.

2.1.3.1 Subject-verb agreement in English and Norwegian lexical verbs

Lexical verbs in English only have overt agreement marking in the third person singular ”-s” (Jensen et al., 2019, 7). The following table (1) shows how the morphological marking in third person singular varies from other conjugations of lexical verbs in the singular in the present tense:

Table 1: Regular conjugation of English lexical verbs in the present tense, represented by ”walk”

	Singular	Plural
1st person	I walk	We walk
2nd person	You walk	You walk
3rd person	He/she/it walk**s**	They walk

In contrast to table 1, there are no overt agreement markers in Norwegian (Jensen et al., 2019, 7), as illustrated by the following table (2):

Table 2: Conjugation of Norwegian lexical verbs in the present tense, represented by ”walk” (gå).

	Singular	Plural
1st person	Jeg går	Vi går
2nd person	Du går	Dere går
3rd person	Han/hun/det går	De går

Looking at the two tables (1, 2), conjugation of lexical verbs in English and Norwegian is similar but not identical. This may be a cause for errors in language learners, which will be discussed further in section 2.2.3.

2.2 Theoretical framework: GenSLA

2.2.1 Generative linguistics

The theoretical framework for language acquisition used in this thesis is generative linguistics, which is an approach that stems from the Universal Grammar (UG) based understanding of language acquisition (Slabakova, 2016, 7). This school of thought sees language acquisition as an innate trait in humans, with UG filling the gaps between what can and cannot be taught through input (Rothman & Slabakova, 2018, 419). More specifically, people are born with mental structures that allow for language acquisition, and these are universal in all humans and allow for the learning of all languages. The rationale behind generative linguistics is a principle called the poverty of the stimulus argument (Garshol, 2019, 11). This principle describes the paradox of why people can produce an infinite amount of grammatical word combinations while only being exposed to a finite number of utterances. Also, children learning the same language are exposed to variable language input. Nevertheless, in most cases, they end up acquiring the same grammar. Accordingly, humans must have a biological predisposition to develop language through linguistic input.

More exact, grammar is seen as a set of principles in generative linguistics, universal to all languages. However, these principles can vary across different parameters (Garshol, 2019, 11). Thus, when children learn a language, they will acquire these grammatical principles differently depending on the parameters present in the given language to which they are exposed. Therefore, input plays a significant role in acquisition even though the structures for acquiring language are present from birth.

2.2.2 Generative linguistics in second language acquisition

GenSLA assumes that L2 acquisition follows the same UG-based rules as L1 acquisition does, based on what White (2003) calls the logical problem (22). The logical problem is, in essence, the same for SLA as what the poverty of the stimulus is for first language acquisition (FLA); the L2 learner acquires subtle grammatical properties that exceed the input they receive, and in this way, L2 learners draw rules from UG. However, some considerable differences must be factored in. One of these differences is the presence of a native language in the mind of the L2

learner, which results in the languages affecting each other. Therefore, understanding the interplay that occurs between all languages is one of the central undertakings of GenSLA research (Rothman & Slabakova, 2018, 419). This interplay can be observed in what has been called the L2 learner's interlanguage – an interim state of a target language (TL) (Hummel, 2014, 65), which can be described as both being constrained by UG (White, 2003, 22) and being affected by the learner's native language. The effects the native language has upon the L2 or TL and vice versa is called crosslinguistic influence (CLI) or transfer. This phenomenon is explained further in the following subsection.

2.2.3 Crosslinguistic influence

Crosslinguistic influence describes how an individual's knowledge of one language affects knowledge of another language (Lipner et al., 2021, 2). As mentioned above, transfer occurs between all languages possessed by a person, not only from the native language toward the L2 or other acquired languages (Ln). In practice, CLI can be observed when a linguistic phenomenon is produced in one of the bilingual's languages that cannot be observed in a monolingual's production of the same language. Furthermore, CLI may be in effect when there are quantitative differences in monolingual and bilingual speakers of a language that can be attributed to another language possessed by the bilingual (Serratrice, 2013, 4). In this sense, differences between a learner's possessed languages may be a source of error, while similarities in these languages may facilitate acquisition.

Much of CLI research focuses on investigating which linguistic domains are more susceptible to transfer, recognizing frequent transfer patterns within specific combinations of languages, and which external and internal factors reinforce or limit CLI. An example is Malcom's (2021) study that investigated CLI in bilinguals speaking Jamaican-creole and Jamaican-English, explicitly focusing on verb-tense, agreement, and copula use, while also exploring how this transfer was affected by factors like motivation and age of arrival (19). Analyzing CLI is made challenging by the overwhelming number of variables that must be accounted for during research; however, some models attempt to illustrate general principles of transfer. These usually discuss whether transfer happens wholesale from one language to another or whether transfer takes effect feature-by-feature, i.e., from one specific linguistic feature to another (Lorenz et al., 2019, 1412).

Transfer can elicit or inhibit target-like language production in the language learner, respectively, known as positive transfer or facilitation and negative transfer or interference. If one assumes that an L2 language learner's L1 is the basis or initial stages of their L2 acquisition – a model proposed in Schwartz & Sprouse (1996) called Full Transfer/Full Access (Schwartz & Sprouse, 1996, 41), these transfer phenomena are bound to occur. According to this model, the learner's interlanguage is constrained by the L1 (Garshol, 2018, 12) and subsequently the source for facilitation or interference in L2 production. Using this model, linguistic phenomena that are dissimilar in the two languages result in interference, while similar linguistic phenomena result in facilitation. CLI can predict some difficulties for L2 learners based on their L1 and L2. There are, however, other factors that play into L2 acquisition. The following two sub-sections will discuss two specific theories explaining why functional morphology is especially hard for L2 learners.

2.2.4 Syntax-before-morphology

Syntax-before-morphology (White, 2003) is a theory arguing that acquisition of syntax drives acquisition of morphology in language learners. A different approach called morphology-before-syntax, arguing that acquisition of morphology drives acquisition of syntax has also been proposed; however, this study adopts the former view.

For a language learner to learn certain linguistic phenomena, like the third person singular "-s", knowledge of several linguistic features must be present. Slabakova (2016, 183) Uses the third person singular ”-s” in an example (1), showcasing the complexities of a tense phrase, which is comprised of morphophonological, semantic and syntactic information:

(1) He often takes the bus

The suffix "-s" carries information about person, tense and aspect. It implicates that the subject is in the third person singular, that the tense is present, and that the aspect is habitual. All these features must be acquired for the learner to use the morpheme reliably correctly; however, that does not mean that these features are obtained at the same time (Slabakova, 2016, 186).

White (2003) argues that L2 English learners acquire abstract knowledge of syntactic structures while still not being stable in their production of inflectional morphology when explaining the Syntax-Before-Morphology theory. This has also been shown in studies. Jensen et al. (2019)

found that L1 Norwegian learners of L2 English had a significantly higher accuracy when judging the grammaticality of sentences by testing knowledge of non-V2 syntax versus knowledge of agreement. In Lardiere (1998a), The test-subject, a Hokkien-Mandarin bilingual learning English, was also found to have far fewer syntactic structure errors than morphological structures, respectively 90 % -100 % correctness and 4.5 % - 34.5 % correctness. This discrepancy indicates that knowledge of syntax is acquired at an earlier stage of language acquisition than functional morphology.

2.2.5 The bottleneck hypothesis

According to The Bottleneck Hypothesis (Slabakova, 2008), functional morphology is challenging to acquire in L2 and Ln acquisition. The hypothesis holds that differences in languages are located in the functional categories, which is also where most language's functional morphology is found (Jensen et al., 2019, 2; Slabakova, 2016, 391). Functional morphology holds much linguistic information and the fact that this aspect is highly varied across languages, makes it a probable source for errors in the language learner. This is in contrast to other linguistic domains, like core syntax, which presumptively is universal in all languages, in the sense that movement obeys the same restrictions (Jensen, 2019, 2). The effects are that functional morphology acts as a bottleneck for L2 production because it affects the acceptability and meaning of the whole sentence (Slabakova, 2021, 321).

Studies on the bottleneck hypothesis have shown that L2 learners do in fact have less accuracy when judging grammaticality related to functional morphology than core syntax. One of these studies is highlighted in Jensen et al. (2019) described earlier in this section. In Slabakova (2019), the author argues that semantics and core syntax are acquired through little effort contra functional morphology, which requires more effort. The author concludes by stating that difficulties in L2 acquisition mainly stem from language architecture and crosslinguistic variation, while frequency, redundancy, and saliency do have an effect, however, it is not crucial.

2.3 Theoretical framework: Instructional second language acquisition

2.3.1 Language instruction – a historical perspective

Explicit grammar instruction has been problematized ever since the naturalistic wave of L2 acquisition gained traction, stating that L2 acquisition followed similar acquisitional patterns to that found in L1 acquisition (Ellis, 2006, 85). The sequences through which L1 learners went when acquiring their first language could, in other words, be detected when observing L2 learners. The natural approach thus became influential, arguing against the use of formal instruction and corrective feedback while preferring the use of methods like comprehensible input. Comprehensible input is based on the view that language learners acquire language when it is understood by the learner (Rodrigo et al., 2003, 54). Corrective feedback, on the other hand, is based on giving explicit, corrective responses on incorrect language production, discussed more closely in 2.3.4. Later evidence has, on the other hand, suggested that instruction and corrective feedback are necessary parts of L2 learning (Spada, 2015, 71). Therefore, the following sub-sections explore implicit language instruction and explicit language instruction.

2.3.2 Implicit and explicit knowledge

The nature of implicit and explicit knowledge and the interplay between these two phenomena are vital when discussing L2 acquisition. The distinction between these two types of knowledge is also significant in language instruction as it inspires teaching methods based on how people process and store information. In theory, understanding these internal mechanisms will help instructors design more effective strategies for language learning. Historically, however, there has been much discussion within the research community, primarily concerning the issue of their dynamic – or rigid – nature, questioning if implicit knowledge can be converted into explicit knowledge and vice versa. In VanPatten (2016), for example, the author claims that explicit knowledge cannot transition into implicit knowledge, or more specifically: "it is difficult to see how explicit knowledge and practice could "guide," "shape," "aid," or in any real sense "support" the growth of implicit knowledge" (655). The following briefly explains what characterizes these two types of knowledge and how the interaction between implicit and explicit knowledge is considered in modern research.

In cognitive psychology, the research field from where the terms explicit and implicit knowledge were coined, the most crucial distinction is whether the knowledge is processed

consciously or unconsciously. In this sense, explicit learning can, for example, take the form of learning the exact grammatical rules of a language. In contrast, implicit learning can take the form of simply being able to understand language input and produce language output without knowing the exact grammar that governs the language. Respectively, the learner can verbalize what they have learned, or they cannot verbalize what they have learned (Ellis et al., 2009, 3). The question of whether these two types of knowledge and learning can transition from one to the other is, in this respect, significant when discussing language learning. It would, for example, answer the question of whether teaching specific grammar will help a language learner towards being fluent in the person's target language, as this would require explicitly taught rules to become automated, implicit knowledge.

Umeda et al. (2019) sought to answer this question through testing L1 Japanese learners of L2 English on their knowledge of English article use. Several acceptability judgement tests were performed, a pre-test and four post-tests, one post-test during the intervention, and three after the intervention (nine 60 minutes lessons over nine weeks). The three post-tests that were performed after the intervention was done one week after the intervention, twelve weeks after post-test 2, and one year after post-test 3. Improvements were found up to post-test 2, however, post-test 4 showed a decrease in retention of the targeted structures, leading the authors to conclude that explicit knowledge of articles is unlikely to be retained after instruction subsides.

Hirakawa, Shibuya & Endo (2019), is a study on two groups of L1 Japanese L2 English students, with a mean age of 19,54 years and 19,15 years. Respectively, the two groups were subjected to explicit instruction and natural exposure. The study sought to compare how the groups responded to these separate types of input when tested on their knowledge of adjective ordering restrictions. The explicit instruction group received 90 minutes of instruction for three weeks, while the natural exposure group were enrolled in three or five-week intense study-abroad programs. The study found that there were no improvements in adjective ordering restrictions in the group receiving natural exposure, while the group that received explicit instruction did show improvement.

2.3.3 The interface models of interaction between explicit and implicit learning

The interface models highlight three different theories on the interplay between explicit and implicit knowledge and try to answer whether conversion between the two is possible. In this sub-section, the interface/non-interface contrast framework is discussed alongside their corresponding pedagogical approaches, focus-on-forms and focus-on-meaning (Dalili, 2011, 2118). According to the non-interface position, implicit knowledge develops through naturally occurring input and interaction, and conversely, implicit knowledge develops through deliberate learning. Primarily, the non-interface-position argues that implicit and explicit knowledge are categorically separate and that the one cannot be converted into the other. (Spada, 2015, 76). The interface position is now often split into two, or rather, the interface position is called the strong interface position, distinguishing it from the related weak interface position (Dalili, 2011, 2118-2119).

Conversely, the non-interface-position claims that implicit and explicit knowledge are not categorically distinct; instead, they are part of a spectrum. Explicit knowledge is converted into implicit knowledge through practice. The weak interface position is more nuanced, arguing that the two can influence and interact with each other, however, not directly convert into each other (Spada, 2015, 76); therefore, the one can affect how the other is acquired or developed. These ways of viewing knowledge and knowledge processing have influenced several methods of language instruction in L2 pedagogy. The first two methods, focus-on-forms and focus-on-meaning, reflect the interface positions in practice (Dalili, 2017, 2118). Focus-on-meaning reflects the non-interface position, and focus-on-forms reflects the strong interface position. The weak interface position is also represented in the focus-on-form instructional approach. Like its counterpart, the weak interface position considers teaching language through explicit and implicit instruction valuable, as both implicit and explicit knowledge can affect the other. The non-interface position and focus on forms framework imply that explicit or declarative knowledge has little effect on L2 instruction. For example, teaching grammar explicitly is unproductive (Spada, 2015, 71-72). On the contrary, the role of explicit, metalinguistic knowledge can help learners be consciously aware of what they learn, which may play an essential part in the language learning process (Toth & Moranski, 2017, 78)

In modern research, it seems that both implicit and explicit knowledge are considered vital factors in L2 acquisition. Metalinguistic knowledge, however, is seemingly relegated to a

supporting role or a "guiding light" for developing implicit knowledge and communicative language skills (Toth & Moranski, 2018, 78). The weak interface position is the view backed by the largest body of evidence (Dalili, 2011, 2120). In this respect, the polarized opinions of the strong- and non-interface-positions seem to have fallen out of favor in modern discourse. This breakthrough is represented in current SLA research (Dalili, 2011, 2121). Accordingly, focus-on-form seems to be the more appropriate framework for L2 instruction. Thus, the weak interface position is adopted in this thesis, and therefore, both a focus on form and meaning are considered to be effective in L2 instruction. The following sub-sections go more into detail on specific instructional technics that follow this framework.

2.3.4 Corrective feedback in ISLA

Corrective feedback in L2 instruction are corrective responses, both explicit and implicit, to an L2 learner's non-target-like production (Li, 2010, 309; Hummel, 2014, 124). Corrective feedback is an essential part of the focus-on-form approach; therefore, figuring out what types of corrective feedback is the most effective. Corrective feedback can be divided into two categories, positive and negative evidence. Positive evidence informs the learner of what correct production sounds or looks like, while negative evidence makes the learner aware of what in their output is incorrect (Li, 2010, 310). In Li (2010), a meta-analysis was conducted to explore different contexts, task types, outcome measures, treatment length, age group and mode of delivery that have been recorded in studies on corrective feedback. There were two factors that were described in the analysis as significant. The first one is that explicit corrective feedback is more effective in the short term. The second one is that implicit corrective feedback was salient long term and that the effects might even increase over time.

2.3.5 Negotiation of meaning

Negotiation of meaning is considered an integral part of task-based language teaching. It is defined in Ellis (2003) as "the process by which two or more interlocutors identify and then attempt to resolve a communication breakdown" (346). It is a teaching method that focuses on communication and is typical for focus-on-meaning based language learning; However, it is also used in focus-on-form based language learning. In an instructional or natural setting, this takes the form of mediation between the teacher and the learner. When input is incomplete or considered unaccepted input, negotiation exchanges take place to make the conversation flow better (Wang, 2019, 83).

Wang (2019) is a data analysis that examined the relevance of creative task engagement, and through this, the author tries to illustrate the importance of meaning negotiation in L2 development. The participants that contributed to the data were 36 English majors (21-22 years). These participants then went through four tasks, two controlled and two creative. The author found that, among adult L2 learners, creative assignments lead to more meaningful negotiation. In particular, were assignments that relied on group cooperation and idea convergence (92).

2.3.6 Language teaching methods used in this study

This study uses two language teaching methods through an intervention in two separate groups to compare their short-term and long-term effects on third-person singular "-s" acquisition. The two language teaching methods are Task Based Language Teaching (TBLT) and Presentation, Practice, Production (PPP). These two methods were chosen because they represent different language learning frameworks. TBLT represents the previously discussed focus on form school of learning, where explicit and implicit learning is used in teaching, however, focusing on implicit learning. On the other hand, PPP represents the focus on forms paradigm of learning, where explicit learning is used. The next two subsections will describe the two teaching methods in more detail.

2.3.7 Task based language teaching

The use of Task based language teaching has seen an increase of interest in the research and educator communities over the last decades (Revesz, 2019, 374). It is an approach to teaching that challenge mainstream ideas, and it has met both resistance and support within academia (East, 2017, 413; Ellis, 2009, 222). TBLT centres around the idea of learning through a focus on meaning, which is achieved using constructions called "tasks". In the TBLT context, a task is a specific definition for a type of instructional activity, and it must follow certain criteria to qualify as a task:

> "1. The primary focus should be on 'meaning' (by which is meant that learners should be mainly concerned with processing the semantic and pragmatic meaning of utterances).
> 2. There should be some kind of 'gap' (i.e. a need to convey information, to express an opinion or to infer meaning).

3. Learners should largely have to rely on their own resources (linguistic and non-linguistic) in order to complete the activity.
4. There is a clearly defined outcome other than the use of language (i.e. the language serves as the means for achieving the outcome, not as an end in its own right)." (Ellis, 2009, 223)

Based on these criteria, teaching grammar using tasks entails not explicitly focusing on the targeted grammatical structures.

Teaching specific grammatical constructions through tasks is done by using what Ellis (2009, 223-224) calls "focused tasks", this is opposed to "unfocused tasks" which are used to teach general language communication. Thus, in a focused task, what is being taught must be undisclosed to the learner to satisfy criteria "1.", as opposed to unfocused tasks where the teaching objective is acknowledged. Thus, a focused task is distinguished from situational grammar exercises because these, in contrast, disclose what the objective grammatical phenomenon is. The task used in this study's intervention is categorized as a focused task and will be explained further in the next section.

2.3.8 Presentation-practice-production

Presentation-practice-production is a teaching method that, as its name suggests, follows three stages that are used throughout a teaching session (Shintani, 2016, 3). The presentation part provides the learners with the linguistic feature that will be this lesson's focus point, which may be done with the goal of making the learners acquire the knowledge inductively or deductively. The second part, practice, makes the learners do exercises with the given language feature, reflecting a high level of teacher control. Checking the learners' production with accuracy is the end goal. Lastly, in the production stage, the learners are given tasks that prompt them to produce the grammatical feature in a less controlled fashion than in the practice part of the lesson. At this stage, the objective is fluency with the linguistic feature through use in autonomous and creative activities (Criado, 2013, 99-100). Thus, in contrast to TBLT, PPP focuses on explicitly teaching the chosen material. Nevertheless, the production phase is similar to an uncontrolled task in the TBLT fashion, with the target linguistic feature being disclosed.

PPP has, in recent years, been criticised for rejecting meaning in favour of forms, especially when keeping in mind that it is based on behaviourist models of teaching. However, some

scholars would still hold that the modern use of PPP still has a place in language teaching as one among several techniques ready for use (Criado, 2013, 111-112). Some scholars also hold that explicit knowledge of forms, or so-called metalinguistic knowledge, may help learners process these features and facilitate development of linguistic competency (Smith & Truscott, 2014, 20) Accordingly, testing the effects of teaching methods like PPP and TBLT should be a priority as the issue is still debated.

2.4 To summarize

Jensen et al. (2019) and Garshol (2018) show that L1 Norwegian learners of L2 English find subject-verb agreement challenging. The findings from these studies are further discussed and compared to the results from the present study (see 5) in section 7.2. Subjet-verb agreement might be challenging for languages learners for several reasons. First, CLI might lead to overgeneralization or negative transfer of certain linguistic features (see 2.2.3). Also, according to white (2003) and the syntax-before-morphology view, knowledge of morphological features is acquired late compared to other linguistic features (see 2.2.4). finally, the bottleneck hypothesis (Slabakova, 2008) argues that variations between languages are located in the functional morphology domain, making this domain a locus of difficulty for language learners (see 2.2.5).

The two instructional approaches used in the present study, PPP and TBLT, represent explicit and implicit learning, respectively. Throughout the history of language instruction, the exclusive use of either explicit or implicit instruction has been scrutinized. Research on modern language instruction seem to favor a more nuanced view, where both explicit and implicit language teaching is applied in instruction – often referred to as a focus on form approach (see 2.3.3). However, the findings in Hirakawa, Shibuya & Endo (2019) and Umeda et al. (2019) contradict each other regarding the effects of explicit grammar instruction. These findings are further discussed in relation to the intervention effects found in the present study in section 7.2.3.

3 Research questions and predictions

This section describes the research questions and subsequent predictions based on this thesis. Section 3.1 overviews the research questions and their rationale, and section 3.2 provides the resulting predictions.

3.1 Research questions

The following RQs will be examined as they include questions that are relevant from both a didactic and a linguistic perspective:

> *RQ 1: Do Norwegian L1 learners of English have little knowledge of the third person singular "-s"?*
>
> *RQ 2: Is there any variation in accuracy on the sub-categories; local singular agreement, local plural agreement, long-distance singular agreement and long-distance plural agreement among L1 Norwegian L2 English learners?*
>
> *RQ 3: Does instruction affect learner's knowledge of the third person singular "-s"?*
>
> *RQ 4: Do different types of instruction affect learner's knowledge of the third person singular "-s" differently?*

Research question 1 is raised because it is of interest to see if the scores from the participants in this study support the data from studies by Jensen (2016), Jensen et al. (2019) and Garshol (2018), that L1 Norwegian learners of English L2 have seemingly disappropriate knowledge of the third person singular "-s" construction. If this is the case, as the above-mentioned studies suggest, then RQs 3 and 4 (and other studies that investigate teaching L1 Norwegians English agreement), are consequential for the Norwegian education system's English instruction.

Research question 2 concerns the sub-conditions in the sentences with subject-verb agreement used in this study. Mainly, whether third person singular "-s" is more challenging to acquire when the subject is plural or when several other lexical items separate the subject and verb. Gathering data on the sub-conditions might highlight what is more or less problematic for L1 Norwegian L2 English learners. Subsequently, this awareness might help to guide instructors trying to teach agreement.

Research question 3 Asks whether there are any changes in accuracy scores between the PPP and TBLT groups in the post-tests. This question is two-sided, as the study tests immediate and long-term effects. Additionally, the two post-tests may reveal discrepancies related to the four sub-conditions. The results may be indicative on whether or not grammar instruction is effective in the short or long-term.

Research question 4 is also concerned with the intervention's short- and long-term effects; however, this question pertains to the contrasts in effect between the two teaching methods. The results may be guiding in when it comes to which teaching method is more effective in teaching grammar. If the results do show any significant differences, the effectiveness of using PPP versus TBLT on teaching agreement should be investigated further.

3.2 Predictions

The RQs will be examined by testing the following predictions:

Theories like the Bottleneck Hypothesis (Slabakova, 2008) and Syntax-before-morphology (White, 2003) hypothesize that functional morphology is hard to acquire for language learners. The literature on functional morphology showcases this difficulty (Slabakova, 2016). Also, data from studies like Jensen (2016), Jensen et al. (2019) and Garshol (2018) support this view. Accordingly, prediction 1 reflects this perspective because the two groups used in this study are comparable to those that were used in the studies mentioned above.

Prediction 1: Subject-verb agreement will be problematic for both groups in the pre-test.

Jensen et al. (2019) showed that the sub-conditions, local plural agreement and long-distance singular agreement, were more challenging to acquire than local singular agreement. Furthermore, the study found that long-distance plural agreement was the most difficult to acquire of the four sub-conditions. The participants of this study were L1 Norwegian L2 English students of 11-12 and 15-18 years and are thus comparable to the participants in this study. Prediction 2 reflects the findings in Jensen et al. (2019).

Prediction 2: The sub-conditions local plural agreement and long-distance singular agreement will be more challenging for the participants than local singular agreement. Long-distance plural agreement will be the most difficult.

In Jensen (2016), there was a correlation between low proficiency test scores and low accuracy on the acceptability judgement task. This data might indicate that some participants, likely those with low proficiency scores, have little knowledge of agreement rules. Therefore, these participants have more room for improvement and have more to gain from instruction.

Prediction 3: Participants with low proficiency test scores will improve more in the post-test and delayed post-test than those with high proficiency test scores.

The groups were subjected to different interventions, one using explicit, forms-focused grammar instruction and one using implicit, form-focused grammar instruction. Umeda et al. (2019) showed immediate improvement through explicit grammar instruction; this was also the case in Hirakawa, Shibuya & Endo (2019). However, the literature favors form-focused instruction for L2 instruction (Dalili, 2011).

Prediction 4: The post-test will show improvement in test scores for both groups.

Prediction 5: Improvements in the PPP group will subside in the delayed post-test, while the improvements will remain in the TBLT group. Alternatively, the improvements in the TBLT group subside less than the improvements in the PPP group.

4 Methodology

The present study uses acceptability judgment tasks with grammatical and ungrammatical sentences and a pedagogical intervention to test L2 English subject verb agreement knowledge in two groups of L1 Norwegian lower secondary school students. The acceptability judgement task in the present study is inspired by Jensen (2016). It uses the same sentences as the author to test for accuracy in judging sentences with subject-verb agreement. By using a quantitative method, such as an acceptability judgement task, numeric data can be gathered statistically and analyzed, with the purpose of determining relationships between the relevant variables (Mackey & Gass, 2005, 137). The study was performed through a pre-test, a post-test (two days after the intervention) and a delayed post-test (two months after the intervention). Having a pre-test and a post-test will ensure the measurement of immediate effects (Mackey & Gass, 2005, 149). Furthermore, the decision to perform the pre-test, intervention and post-test in close succession (four days from pre-test to post-test) was made to ensure that the results from the post-test were immediate. However, a point can be made that the post-test should have been performed immediately, in its literal sense, to eliminate variables like sleep quality and other factors that might make individual differences in the participant's retention and post-test scores. Improvements and limitations will be discussed further in section 5.

As this study compares the effects of L2 instructional methods, measuring long-term effects is essential. A delayed post-test in addition to the post-test was used for this purpose. One drawback of using a delayed post-test is that extra-experimental exposure will be more significant, and, additionally, there is also a risk of losing participants (Mackey & Gass, 2005, 149). Nevertheless, using test scores from a delayed post-test is vital for determining the long-term effects of the intervention in this study.

The groups that were used are not randomly assigned; hence, the intervention is a quasi-experiment. (Mackey & Gass, 2005, 146). The quasi-experiment was performed through a classroom intervention. This means that the test subjects are exposed to manipulated input to measure the effects of the chosen teaching methods, TBLT and PPP. The core of a quasi-experiment design is the intervention itself. Therefore, the intervention will take place after the pre-test to ensure that the participant's test scores in the post-test and delayed post-test have a baseline that can be used to measure the effects of the intervention.

Increased control is one of the strengths of using an intervention, as it makes it easier to control variables and tailor tasks to the relevant grammatical phenomenon (Loewen & Philip, 2012, 61). Weaknesses of using an intervention-based approach include the artificial nature of manipulating input and other variables, and mapping and controlling all variables in a classroom is impossible (Loewen & Philip, 2012, 61-62). However, the teaching methods used in this experiment (TBLT and PPP) are commonly used in classroom instruction. This similarity may mask the artificial aspect of the experiment to a certain degree, hopefully making the intervention seem more like a natural part of the participant's education. Also, having the intervention take place in a naturalistic context (elaborated upon in section 6.4) like a lower secondary school classroom, with intact classes, may strengthen the face validity of this study's findings because it may genuinely reflect authentic settings that L1 Norwegian L2 English learners find themselves in throughout their education (Mackey & Gass, 2005, 143).

The pedagogical intervention uses one contrasting teaching session for each of the two groups (the choice of which class underwent which intervention was arbitrarily chosen). The goal is to test the effects of the distinctive teaching sessions in the post-test and delayed post-test and compare them to the results from the pre-test. Group 1 was subjected to a PPP-based teaching session, and group 2 was subjected to a TBLT-based teaching session. The groups also went through a sub-set of the Standardized Oxford Proficiency Test to check their general English proficiency level and a language background questionnaire. All information surrounding the tests and intervention was given in Norwegian to avoid any misunderstandings. The participants were not explicitly informed that they were being tested on their knowledge of English subject-verb agreement. Further descriptions of the intervention are presented in 4.3.

4.1 Participants

All participants used in this study gave their informed consent (see appendix 1) to participate in this study. Approval was given through correspondence with NSD (see appendix 2), and measures have been taken to ensure the participants anonymity.

The participants in the study are L1 Norwegian lower secondary school students. The participants were going through their final year of lower secondary school. The study was conducted between early December and into early February. The participants were between 14 and 16 years old and enrolled in 10th grade. The intervention used intact groups where the

pupils were assigned to group conditions based on their respective classes. Accordingly, the two groups were parallel school classes, going through the same curriculum simultaneously. Group 1 consists of 16 participants (n = 16), 6 females and 10 males, and group 2 consists of 20 participants (n = 20), 8 females and 12 males. The PPP-group had a mean age of acquisition (AoA) of 6.076 and the TBLT group had mean AoA of 6,73.

Due to logistical difficulties, not every participant went through all three of the acceptability judgement tasks. In group 1, 16 participants participated in the pre-test, 11 in the post-test and 11 in the delayed post-test. In group 2, 20 participants participated in the pre-test, 15 in the post-test and 13 in the delayed post-test. In fact, only half (p = 8) of group 1, the PPP group, were measured at all points, and only (p = 11) out of group 2, the TBLT group were measured at all points. Also, 1 participant in group 1 and 1 participant in group 2 did not have Norwegian as their L1, meaning that their data is not included in the results. Therefore, the actual number of participants used is 15 for the PPP-group (n=15) and 19 for the TBLT-group (n=19). The following table (3) is an overview of the participants that have contributed to the dataset in this study:

Table 3: An overview of the participants that have contributed to the study's dataset

	PPP-group (*n*=15)	TBLT-group (*n*=19)
Birthyear	2006	2006
Mean AoA	6,076	6,73
L1	Norwegian	Norwegian

4.2 Tests

A total of four tests were conducted: The pre-test, proficiency test and post-test were performed in the same week. First, the pre-test (acceptability judgement test) and the proficiency test were conducted during the same school hour (one school hour = 60 minutes, approximately 20 minutes, separately, for both the proficiency test and the acceptability judgement task). Then, the following school hour was used to conduct the intervention. The third school hour was used for the post-test, two days after the intervention for both groups. The delayed post-test was conducted two months after the intervention.

The Oxford Proficiency Test is described in section 4.2.1, and the acceptability judgement task is explained in 4.2.2, which also showcases some of the sentences used in the acceptability

judgment task, with examples from the four sub-conditions tested for; local singular, local plural, long-distance singular, and long-distance plural agreement.

4.2.1 Language background, questionnaire and Oxford Proficiency Test

4.2.1.1 Language background questionnaire

The language background questionnaire asked the participants about gender, birth year, what language they used at home and when the participant entered the Norwegian school system, first and foremost to map their age-group. The participants age of onset of acquisition (AoA) and home language also provide additional information to be used in the discussion. The participants were also asked what other languages they have learned/are learning, if they use other languages in their daily lives, what they use them for, and, finally, what year the participant started learning English. This was primarily done to check for age and language discrepancies and isolate participants who do not have Norwegian as their first language. The questionnaire can be viewed in appendix 8.

4.2.1.2 The standard Oxford Proficiency test

A subset of a standard Oxford Proficiency test was used to check the participant's general knowledge of English (Appendix 1.) so that it may be compared to the participants accuracy scores in the acceptability judgement task pre-test (see RQ 1). The same proficiency test has been used in other studies alongside acceptability judgement tests (Jensen, 2016; 2019). The test is a multiple-choice task where the goal is for the participants to mark the word that fits in the corresponding sentence to make it grammatically correct. See example (2)

(2) Very _____ people can travel abroad.

☐ less

☐ little

☐ few

There was a total of 40 multiple-choice questions following the same formula in the proficiency test, including example (2). The test is split into two parts, with the latter 20 questions forming a continuous story. The rest of the proficiency test can be viewed in appendix 3.

4.2.2 Acceptability judgement tasks

The test used to measure the participant's knowledge of the third person singular "-s" was an acceptability judgement task. This method was chosen because it makes it possible to systematize quantitatively which sentences the participants find acceptable as grammatical or ungrammatical. In an acceptability judgement task, the participants judge the grammaticality of a presented sentence (Mackey & Gass, 2005, 49) or, more precisely, their acceptance of the sentence as grammatical. The answers were scored in a binary system as grammatical or ungrammatical. The score sheet given to the participants is shown in figure 1. There is a total of 48 sentences, 24 of which relate to subject-verb agreement; the rest were fillers (see appendix 5 and 6). Fillers were used to draw the participant's attention away from the actual sentences they were being tested on to prevent them from realizing the test's purpose. The sentences with subject-verb agreement were divided into four categories, representing the four sub-conditions. There were six sentences per category:

1) Main clauses with local agreement, singular subjects

2) Main clauses with local agreement, plural subjects

3) Main clauses with long distance agreement, singular subjects

4) Main clauses with long distance agreement, plural subjects

Table 4 shows an example from each of the four sub-conditions:

Table 4: An example from each sub-condition, both grammatical and ungrammatical

Category:	**Sentence**
Local singular agreement	The student loves to read books about football
	*The student love to read books about football
Local plural agreement	The teachers give their students a lot of homework
	*The teachers gives their students a lot of homework
Long-distance singular agreement	The girl with golden earrings takes the bus to school
	*The girl with golden earrings take the bus to school
Long-distance plural agreement	The parents with the nice car talk to their kids
	*The parents with the nice car talks to their kids.

Table 4 shows the words used in the sentences, which were taken from a word frequency list (Jensen, 2016). The sentences also follow rules that account for extragrammatical factors like length, lexical content and processing difficulty (Dabrowska, 2010, 5) by keeping the sentences short (8-10 words) and the words familiar.

4.2.2.1 Acceptability judgement task: Sequence of events

The test was presented to the participants using a PowerPoint presentation. Each participant was given a scoring sheet and told approximately how long (20 minutes) the test would take. The presentation itself had instructions and examples, showing the participants how and when to use their scoring sheet. The scoring sheet had a number corresponding to each sentence, as can be seen in figure 1. The entire sheet can be viewed in appendix 4. The students had ten seconds per sentence to judge whether they found it grammatically acceptable or unacceptable. Ten seconds was chosen as a time limit because the task's purpose is to check for knowledge of the subject's grammar, i.e., what would be used in their communication, as opposed to whether the person remembers the formal rule and can rationalize their way to an answer (Mackey & Gass, 2005, 51). Every sentence was read aloud by the instructor so that the participants could decide based on two modalities.

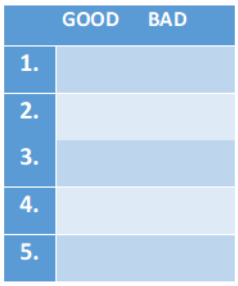


Figure 1: Showing five of the 48 identical rows in the original sheet

Each of the 48 sentences had an ungrammatical version to choose from to ensure that there were an equal number of grammatical and ungrammatical sentences. Because the test was performed three times, as a pre-test, post-test and delayed post-test, the order and grammaticality of the sentences had to be changed for each test so that none of the students could anticipate the correct judgments. The changes were done pseudo-randomly.

4.3 The intervention

In this sub-section, the course of the interventions is explained. Firstly, the PPP instruction in 4.3.1 and then the TBLT instruction in 4.3.2. The following tables (5) and (6) are overviews of the two sessions.

Table 5: An overview of the PPP instructional session

Activity	**Approximated time span**
Part 1: Presentation	15 minutes
Part 2: Practice	20 minutes
Part 3: Production	20 minutes

Table 6: An overview of the TBLT instructional session

Activity	**Approximated time span**
Priming activity	5 minutes
Main task	40 minutes
Post-task	10 minutes

4.3.1 Intervention group 1: PPP

This session lasted 60 minutes. The first five minutes were spent on getting started, waiting for all the students to arrive and noting absentees, leaving a total of 55 minutes for the planned teaching session, which was split into three parts.

Part 1: Presentation

The students were presented with examples and explanations of the target grammatical structure through a slide show (examples of the slides are shown in figures 2 and 3). Next, they were given contexts and "everyday use" examples and were asked about the general rule of the structure, i.e., what does agreement mean in a grammar context. Students had opportunities to learn deductively (see 2.3.8), as examples of the general rules were presented before the instructor explained the rules explicitly. This part took approximately 15 minutes.

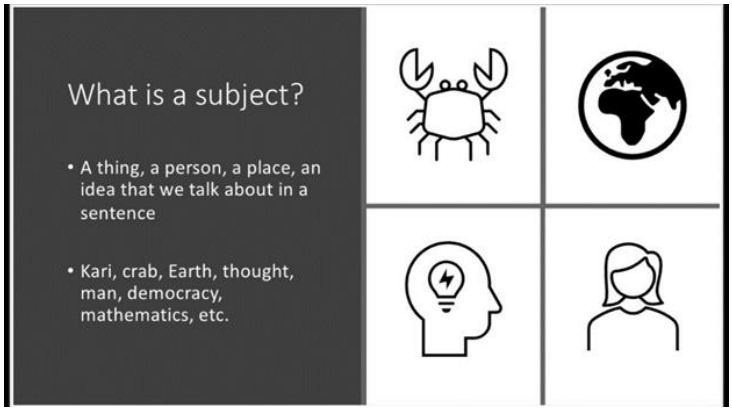


Figure 2: Screenshot from PowerPoint lecture

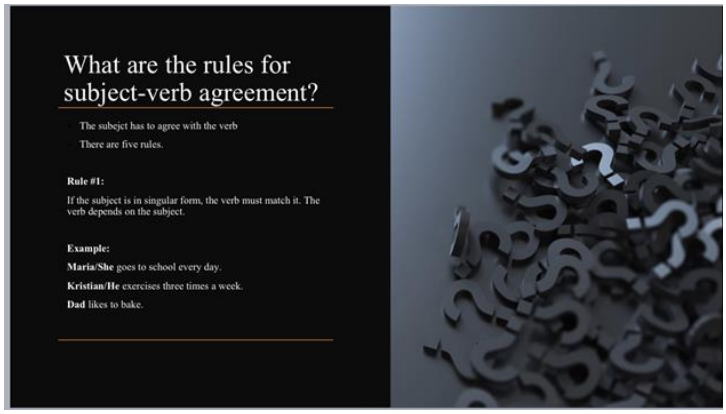


Figure 3: Screenshot from PowerPoint lecture

Part 2: Practice

The Practice part of the session is usually a teacher-controlled activity that seeks to raise the student's fluency in the targeted structure to be used more freely in the production stage (Criado, 2013, 99) (see 2.3.8). In this part, the students were first given situational grammar exercises that targeted both long and short distance agreement. Then, the instructor went through the answers in a plenary session.

Afterwards, the students were divided into pairs and given bits of paper with text (see appendix 7). These bits of paper had a subject or a verb written on them. The task is for the pairs to put together subjects and verbs that agree, where the couple that finishes the task correctly, the fastest, wins a small price. These two activities combined took approximately 20 minutes.

Part 3: Production

A "production activity" should make the students use the targeted construction more freely and creatively than in the practice part of the session (Criado, 2013, 100) (see 2.3.8). In the last part of the session, the students were divided into new groups of three. Here, all groups were given three pictures, each member a different one. Firstly, they wrote down descriptions of what actions are happening in the picture - in at least five sentences. When everyone was done writing down their descriptions, they discussed their descriptions of the pictures and commented on aspects the others in the group might have missed. This activity was designed to elicit production of sentences containing subject-verb agreement. The activity took approximately 20 minutes.

4.3.2 Intervention group 2: TBLT

Similarly to the PPP class, this session lasted 60 minutes. The first five minutes were spent on getting started, waiting for all the students to arrive and noting absentees, leaving 55 minutes for the planned teaching session, which was also split into three parts.

Part 1: priming activity

The session started with a priming activity. Priming activities are not necessary for the activity to qualify as a task, and it is not supposed to touch upon the actual teaching objective. Instead, the priming activity is supposed to create expectation and arouse the students by, for example, telling them how the task may relate to their real life. The goal is to elicit motivation and counteract negative expectations in the students. Other functions of a pre-task include preparing

the students for the task ahead, allowing them to be adequately equipped with knowledge about the task and resources for completion (Ellis et al., 2019, 210).

The group and the instructor discussed the concept of "philanthropy". The instructor asked the students what they associated with philanthropy. Then, the students and the instructor concluded what the concept meant together. This activity took five minutes.

Part 2: Main task

During the main task, where students were prompted to produce sentences containing the third person singular "-s", corrective feedback (CF) (see 2.3.4) was used to guide the students' output during the task completion. This type of correction is called "reactive focus on form, " which mainly provides feedback on errors during task performance (Ellis et al., 2019, 222-223). In other words, the participants performed the task in written form while the instructor assessed their production (see 2.3.7) and gave CF on the targeted grammatical feature.

The students were divided into groups of three. Each member of a group chose a celebrity that they were going to research. The theme was "philanthropy", and the goal was for the students to write a short essay (1-2 small paragraphs at about 5-10 lines each) on what good their chosen celebrity does for society. When they had completed their essay, they read and revised each other's essays. Lastly, the students summed up their essays for the class, which segued into the post-task. The participants were given 40 minutes for this task.

Part 3: Post-task

finally, through a post-task, the target grammatical structure was addressed. This activity contained a plenary repetition of "observed" grammatical errors during their task completion. The instructor told the students what grammatical errors seemed to be recurring in their texts and focused on the third person singular "*-s*" in a plenary session. Other errors were also addressed so as not to disclose the tasks' purpose. This part took approximately ten minutes.

5 Results

This section will showcase the results from the proficiency test and the acceptability judgement tasks. First, 5.1 displays the proficiency test scores, and then the sub-conditions from each acceptability judgement task are displayed in consecutive order in 5.2. Next, 5.3 shows the correlation between the proficiency test scores and the acceptability judgement task pre-test, post-test and delayed post-test. Section 5.4 compares the overall results from the acceptability judgment tasks related to the interventions.

5.1 Standardized Oxford Proficiency test

A standardized Oxford proficiency test containing 40 multiple-choice questions was used. The participant must choose an answer based on what word is grammatically correct in the given sentence (see example 2). The test scores ranged from 17 to 36 in the PPP group and from 16 to 35 in the TBLT group. The results for the PPP and TBLT groups are displayed in table 7. One participant from the PPP group and two participants from the TBLT group did not attend the proficiency test. Also, the results from the two participants with other L1s than Norwegian have been excluded.

Table 7: Oxford proficiency test scores from both groups

	PPP scores	TBLT scores
	17	16
	17	21
	28	25
	28	25
	29	28
	30	29
	30	30
	30	31
	32	32
	33	33
	33	34
	34	34
	36	34
	36	34
		35
		35
		36
Mean	29,5	30,1176

Table 7 shows that the mean score from the PPP group was 29,5, and the mean score from the TBLT group was 30,1176. Thus, with a mean difference of 0,6176, the groups are comparable. These results are linked to the acceptability judgement task scores in 5.3.

5.2 Acceptability judgement task scores

The acceptability judgment task consists of 48 sentences, 24 of which are sentences that relate du subject-verb agreement. Four sub-conditions are represented among these 24 sentences: Local singular agreement, local plural agreement, long-distance singular agreement and long-distance plural agreement. The rest of the sentences are fillers. In this sub-section, the acceptability judgement task test scores are displayed in consecutive order, detailing each sub-condition. The results from the tests are presented in the tables below. The tables containing the scores from the four sub-conditions show the number of correct answers per sentence and accuracy percentage. The accuracy percentage display how many per cent of the total answers were correct per sub-condition.

5.2.1 Pre-test

The pre-test scores from both groups are shown in table 8. A closer analysis is displayed in tables 9, 10, 11 and 12, showing the scores in the four sub-conditions.

Table 8: Pre-test scores from the PPP group and the TBLT group

Pre-test	PPP group	TBLT group
	10	12
	11	12
	12	12
	12	13
	13	13
	13	14
	13	14
	13	14
	13	15
	13	15
	14	16
	16	16
	16	16
	20	16
	21	16
		19
		19
		19
		21
Mean score	14	15,3684211

Table 8 shows the pre-test scores from both groups. In the pre-test, there were 15 participants (P=15) from the PPP group and 19 participants (P=19) from the TBLT group. The pre-test scores show an average of 14 (58,33 %) in the PPP group and 15,368 (64,47 %) in the TBLT group, revealing a difference of 1,368. According to Brown (1974), 90 % accuracy can be categorized as target-like. Consequently, both groups seem to find subject-verb agreement challenging, per **prediction 1**. The following, table 9 shows the scores from the sub-condition targeting local singular agreement.

Table 9: Pre-test scores from both groups in the local singular agreement sub-condition

	Sum of Loc.sing 1	Sum of Loc.sing 2	Sum of Loc.sing 3	Sum of Loc.sing 4	Sum of Loc.sing 5	Sum of Loc.sing 6	Sum total	Aaccuracy percentage
Pre-test	**16**	**31**	**17**	**34**	**10**	**33**		
PPP	4	13	8	15	4	14	58	64,44 %
TBLT	12	18	9	19	6	19	83	72,81 %
Totalsum	**16**	**31**	**17**	**34**	**10**	**33**		

When examining the scores on the sentences related to the local singular agreement sub-condition in table 9, the data reveals an accuracy percentage of 64,44 for the PPP group and 72,81 for the TBLT-group, a difference of 8,37 %. The table also shows the participant's accuracy in individual sentences. For example, 4 of the 15 participants in the PPP group answered correctly on the local singular agreement sentence 1. Next, table 10 shows the scores from the sub-condition targeting local plural agreement.

Table 10: Pre-test scores from both groups in the local plural agreement sub-condition

	Sum of Loc.plu 1	Sum of Loc.plu 2	Sum of Loc.plu 3	Sum of Loc.plu 4	Sum of Loc.plu 5	Sum of Loc.plu 6	Sum total	Accuracy percentage
Pre-test	**6**	**33**	**14**	**30**	**12**	**27**		
PPP	1	14	5	14	4	13	51	56,67 %
TBLT	5	19	9	16	8	14	71	62,28 %
Totalsum	**6**	**33**	**14**	**30**	**12**	**27**		

Table 10 shows that the scores on the sentences related to local plural agreement reveal an accuracy percentage of 56,67 for the PPP group and 62,28 for the TBLT-group, a difference of 5,61 %. The data also displays that accuracy on local plural agreement trials was lower for both groups (7,77 % for the PPP group and 10,53 % for the TBLT group) than on trials targeting local singular agreement. The next table (11), displays the results from the sub-condition targeting long-distance singular agreement.

Table 11: Pre-test scores from both groups in the long-distance singular agreement sub-condition

	Sum of Long.sing 1	Sum of Long.sing 2	Sum of Long.sing 3	Sum of Long.sing 4	Sum of Long.sing 5	Sum of Long.sing 6	Sum total	Accuracy percentage
Pre-test	**17**	**33**	**20**	**29**	**20**	**27**		
PPP	7	15	5	12	6	11	56	62,20 %
TBLT	10	18	15	17	14	16	90	78,95 %
Totalsum	**17**	**33**	**20**	**29**	**20**	**27**		

Table 11 shows that the scores on the sentences related to long-distance singular agreement reveal an accuracy percentage of 62,20 for the PPP group and 78,95 for the TBLT group. This variance makes up a difference of 16,75 %, which is the greatest among the four sub-conditions in the pre-test. In this sub-condition, both groups have higher accuracy than in the local plural sub-condition, with an increase of 5,54 % and 16,67 % in the PPP and TBLT groups, respectively. The following table (12) shows the scores from the sub-condition targeting long-distance plural agreement.

Table 12: Pre-test scores from both groups in the long-distance plural agreement sub-condition

	Sum of Long.plu 1	Sum of Long.plu 2	Sum of Long.plu 3	Sum of Long.plu 4	Sum of Long.plu 5	Sum of Long.plu 6	Sum total	Accuracy percentage
Pre-test	**12**	**21**	**13**	**18**	**16**	**13**		
PPP	8	10	6	11	5	5	45	50 %
TBLT	4	11	7	7	11	8	48	42,11 %
Totalsum	**12**	**21**	**13**	**18**	**16**	**13**		

Table 12 shows that the scores on the sentences related to long-distance plural agreement reveal an accuracy percentage of 50 for the PPP group and 42,11 for the TBLT-group, which makes a difference of 7,89 %. Accuracy decreases for both groups from the long-distance singular sub-condition to the long-distance plural sub-condition by 12,20 % and 36,84 % for the PPP and TBLT groups, the most significant difference among all categories in the pre-test. Long-distance plural is the only sub-condition in the pre-test where the PPP group outperforms the TBLT-group.

5.2.1.1 Pre-test summary

The results are in line with **prediction 2** (see 3.2). The pre-test scores reveal that local singular agreement was the least challenging sub-condition for the PPP group and the second least challenging sub-condition for the TBLT group. However, the differences between the three sub-conditions: Local singular, local plural and long-distance singular agreement were not substantial. Thus, no conclusions can be drawn based on the results from these three sub-conditions in the pre-test. Conversely, the sub-condition that tested for long-distance plural agreement was clearly the most challenging, as per prediction.

5.2.2 Post-test

The post-test scores from both groups are shown in table 13. A closer analysis is displayed in tables 14, 15, 16 and 17, showing the scores in the four sub-conditions.

Table 13: Post-test scores from the PPP group and the TBLT group

Post-test	PPP group	TBLT group
	12	10
	13	10
	13	10
	14	11
	15	12
	15	12
	15	12
	15	13
	16	13
	18	13
	22	15
		15
		18
		18
Mean score	15,2727273	13

In the post-test, there were 11 participants (P=11) from the PPP group and 15 participants (P=15) from the TBLT group. The post-test scores show an average score of 15,27 in the PPP group and 13 in the TBLT group, revealing a difference of 2,27. The following are the tables displaying the results from the four sub-conditions. First, table 14 shows the results from the sub-condition targeting local agreement.

Table 14: Post-test scores from both groups in the local singular agreement sub-condition

	Sum of Loc.sing 1	Sum of Loc.sing 2	Sum of Loc.sing 3	Sum of Loc.sing 4	Sum of Loc.sing 5	Sum of Loc.sing 6	Sum total	Aaccuracy percentage
Post-test	24	15	25	9	21	9		
PPP	11	6	11	4	9	5	46	69,70 %
TBLT	13	9	14	5	12	4	57	63,33 %
Totalsum	24	15	25	9	21	9		

Table 14 shows that the accuracy percentage of the PPP-group in the local singular agreement sub-condition in the post-test was 69,70, while the accuracy percentage of the TBLT group was 63,33, revealing a difference of 6,37 %. As opposed to the pre-test, the PPP group had the highest accuracy percentage in the local singular sub-condition in the post-test. The following table (15) displays the scores from the sub-condition targeting local plural agreement.

Table 15: Post-test scores from both groups in the local plural agreement sub-condition

	Summer av Loc.plu 1	Summer av Loc.plu 2	Summer av Loc.plu 3	Summer av Loc.plu 4	Summer av Loc.plu 5	Summer av Loc.plu 6	Sum total	Aaccuracy percentage
Post-test	22	10	17	7	19	11		
PPP	11	5	10	3	10	6	45	68,18 %
TBLT	11	5	7	4	9	5	41	45,56 %
Totalsum	22	10	17	7	19	11		

Table 15 shows that the accuracy percentage of the PPP-group in the local plural agreement sub-condition in the post-test was 68,18, while the accuracy percentage of the TBLT group was 45,56, revealing a difference of 22,62 %. The variance in accuracy from the local singular sub-

condition is at 1,52 % and 17,77 % for the PPP group and TBLT group, respectively. Again, as opposed to the pre-test, the PPP group outperforms the TBLT group in this sub-condition. Next, table 16 shows the results from the sub-condition targeting long-distance singular agreement.

Table 16: Post-test scores from both groups in the long-distance singular agreement sub-condition

	Summer av Long.sing 1	Summer av Long.sing 2	Summer av Long.sing 3	Summer av Long.sing 4	Summer av Long.sing 5	Summer av Long.sing 6	Sum total	Aaccuracy percentage
Post-test	**23**	**11**	**20**	**6**	**22**	**10**		
PPP	11	4	9	3	10	2	39	59,09 %
TBLT	12	7	11	3	12	8	53	58,89 %
Totalsum	**23**	**11**	**20**	**6**	**22**	**10**		

Table 16 shows that the accuracy percentage of the PPP-group in the long-distance singular agreement sub-condition in the post-test was 59,09, while the accuracy percentage of the TBLT-group was 58,89, revealing a difference of 0,2 %. The difference from the local plural sub-condition is 9,09 % and 13,33 % for the PPP and TBLT groups, respectively. Like in the pre-test, the TBLT group performed better at the long-distance singular sub-condition than in the local plural sub-condition. The trend also continues, with the PPP group outperforming the TBLT group in this condition, unlike in the pre-test. The next table (17) displays the scores from the sub-condition targeting long-distance plural agreement.

Table 17: Post-test scores from both groups in the long-distance plural agreement sub-condition

	Summer av Long.plu 1	Summer av Long.plu 2	Summer av Long.plu 3	Summer av Long.plu 4	Summer av Long.plu 5	Summer av Long.plu 6	Sum total	Aaccuracy percentage
Post-test	**14**	**6**	**19**	**7**	**11**	**12**		
PPP	7	4	11	2	8	6	38	57,58 %
TBLT	7	2	8	5	3	6	31	34,44 %
Totalsum	**14**	**6**	**19**	**7**	**11**	**12**		

Table 17 shows that the accuracy percentage of the PPP-group in the long-distance plural agreement sub-condition in the post-test was 57,58, while the accuracy percentage of the TBLT-group was 34,44, revealing a difference of 23,14 %. The variance from the long-distance singular sub-condition is 1,51 % and 24,45 % for the PPP group and the TBLT group, respectively. Similarly to the pre-test results, the PPP group find the Long-distance plural condition less challenging than the TBLT group.

5.2.2.1 Post-test: Summary

Both groups found the local singular sub-condition the least challenging in the post-test, and both groups found the long-distance plural sub-condition to be the most challenging. Accordingly, the results from the post-test follow **prediction 2**, as opposed to the results from the pre-test. Unlike the results from the pre-test, the PPP group had the highest accuracy percentage in all sub-conditions. One explanation might be the short-term effects of the intervention. On the other hand, both groups had fewer participating students; thus, it is

impossible to interpret the results as representative of the group's knowledge of English agreement making.

5.2.3 Delayed post-test

The post-test scores from both groups are shown in table 18. A closer analysis is displayed in tables 19, 20, 21 and 22, showing the scores in the four sub-conditions.

Table 18: Delayed post-test scores from the PPP group and the TBLT group

Delayed post-test	PPP group	TBLT group
	9	10
	10	12
	11	12
	12	13
	16	13
	16	14
	16	14
	16	15
	19	20
	21	21
		22
Mean score	14,6	15,0909091

As shown in table 18, There were 10 participants (P=10) from the PPP group and 11 participants (P=11) from the TBLT group in the delayed post-test. The delayed post-test scores reveal an average of 14,6 in the PPP group and 15,091 in the TBLT group. The following tables show the results from the sub-conditions in the delayed post-test. First, table 19 displays the results from the sub-condition targeting local singular agreement.

Table 19: Delayed post-test scores from both groups in the local singular agreement sub-condition

	Sum of Loc.sing 1	Summer av Loc.sin	Sum of Loc.sing 3	Sum of Loc.sing 4	Sum of Loc.sing 5	Sum of Loc.sing 6	Sum total	Aaccuracy percentage
Delayed post-test	11	8	7	20	19	20		
PPP	4	4	2	10	8	9	37	61,70 %
TBLT	7	4	5	10	11	11	48	72,73 %
Totalsum	11	8	7	20	19	20		

Table 19 shows that the accuracy percentage of the PPP-group in the local singular agreement sub-condition in the delayed post-test was 61,70, while the accuracy percentage of the TBLT-group was 72,73, revealing a difference of 11,03 %. The TBLT group has the highest accuracy percentage in this sub-condition, like in the pre-test. Table 20 shows the results from the sub-condition targeting local plural agreement.

Table 20: Delayed post-test scores from both groups in the local plural agreement sub-condition

	Sum of Loc.plu 1	Sum of Loc.plu 2	Sum of Loc.plu 3	Sum of Loc.plu 4	Sum of Loc.plu 5	Sum of Loc.plu 6	Sum total	Aaccuracy percentage
Delayed post-test	**6**	**21**	**12**	**8**	**17**	**18**		
PPP	2	10	7	5	8	7	39	65,00 %
TBLT	4	11	5	3	9	11	43	65,15 %
Totalsum	**6**	**21**	**12**	**8**	**17**	**18**		

Table 20 shows that the accuracy percentage of the PPP-group in the local plural agreement sub-condition in the delayed post-test was 65,00, while the accuracy percentage of the TBLT-group was 65,15, revealing a difference of 0,15 %. The variance from the local singular sub-condition is 3,3 % and 7,58 % for the PPP group and the TBLT group, respectively; however, the former group has a higher accuracy percentage than in the local singular sub-condition, as opposed to the latter group. Similarly to the pre-test results, the TBLT group find the local plural condition less challenging than the PPP group. Next, table 21 displays the scores from the sub-condition targeting long-distance singular agreement.

Table 21: Delayed post-test scores from both groups in the long-distance singular agreement sub-condition

	Sum of Long.sing 1	Sum of Long.sing 2	Sum of Long.sing 3	Sum of Long.sing 4	Sum of Long.sing 5	Sum of Long.sing 6	Sum total	Aaccuracy percentage
Delayed post-test	**9**	**18**	**10**	**16**	**14**	**19**		
PPP	5	8	5	7	8	9	42	70,00 %
TBLT	4	10	5	9	6	10	44	66,67 %
Totalsum	**9**	**18**	**10**	**16**	**14**	**19**		

Table 21 shows that the accuracy percentage of the PPP-group in the local plural agreement sub-condition in the delayed post-test was 70,00, while the accuracy percentage of the TBLT-group was 66,67, revealing a difference of 3,33 %. Both groups have higher scores than in the local plural sub-condition, with a difference of 5 % and 1,52 % for the PPP group and the TBLT group, respectively. Similarly to the pre-test, both groups find long-distance singular agreement easier than local plural agreement. The last table (22) displaying the sub-conditions from the acceptability judgement tasks shows the sub-condition targeting long-distance plural agreement.

Table 22: Delayed post-test scores from both groups in the long-distance plural agreement sub-condition

	Sum of Long.plu 1	Sum of Long.plu 2	Sum of Long.plu 3	Sum of Long.plu 4	Sum of Long.plu 5	Sum of Long.plu 6	Sum total	Aaccuracy percentage
Delayed post-test	**9**	**13**	**4**	**14**	**10**	**14**		
PPP	5	6	1	7	3	6	28	46,67 %
TBLT	4	7	3	7	7	8	36	54,55 %
Totalsum	**9**	**13**	**4**	**14**	**10**	**14**		

Table 22 shows that the accuracy percentage of the PPP-group in the local plural agreement sub-condition in the delayed post-test was 46,67, while the accuracy percentage of the TBLT-group was 54,55, revealing a difference of 7,88 %. The data reveals an accuracy discrepancy between the long-distance singular and the long-distance plural sub-conditions of 23,33 % and

12,12 % for the PPP and TBLT groups, respectively. Like in the pre-test and post-test, the long-distance plural sub-condition appears to be the most challenging for both groups.

5.2.3.1 Delayed post-test: Summary

Analogously with the post-test results, both groups found the long-distance plural sub-condition to be the most challenging in the delayed post-test, following **prediction 2**. On the other hand, the PPP group found the local plural and long-distance singular sub-conditions less challenging than the local singular sub-condition, contradicting **prediction 2**. The TBLT group had the highest accuracy percentage in all sub-conditions except for the local plural sub-condition, unlike the results from the pre-test and post-test.

Only one of the groups improved its accuracy percentage from the pre-test overall. The scores went from 14 in the PPP group and 15,368 in the TBLT group to a mean score of 14,6 in the PPP group and 15,091 in the TBLT group. However, as fewer of the students participated in the delayed post-test than in the pre-test, it is impossible to assume that this improvement is a long-term effect of the intervention.

5.2.4 Summary: All tests

When examining the test results, there is one common observation in all three tests: Long-distance plural agreement is the most challenging sub-condition for both groups. The accuracy percentage in the other sub-conditions varied greatly throughout the three tests, so no other common pattern seemed evident. This may be caused by the variability in number of students participating in each test, as the parts of the groups that participated in the post-test and delayed post-test are not representative of the group's total number. Additionally, the group sizes were initially small, which may further explain the results. Despite this, the accuracy percentage in the long-distance plural sub-condition is relatively low compared to the other sub-conditions in all tests. If one considers the participants of each test as isolated groups, all six groups find long-distance plural agreement challenging. Thus, we can conclude that the second part of **prediction 2** holds true for these groups.

5.3 Comparing proficiency test and acceptability judgement task scores.

This sub-section compares the participant's proficiency score and accuracy percentage in the acceptability judgement tests. The results from the PPP group are discussed in 4.3.1, and the results from the TBLT group are discussed in 4.3.2. Table 23 is an overview of the results from

the PPP group, and table 24 presents the TBLT group's results. In 4.3.3, **prediction 3**, which stated that participants with low proficiency scores would improve more after the intervention than the participants with high proficiency scores, is discussed.

5.3.1 PPP group results

The following table (23) shows the participant's proficiency test scores on the left, with corresponding acceptability judgement task accuracy percentage on each test in the subsequent columns. The last two columns display differences in accuracy percentage from the pre-test to the post-test and delayed post-test. Empty cells represent non-participation.

Table 23: Data from the PPP group's proficiency test and acceptability judgement tasks

Proficiency test scores	Pre-test correctness	Post-test correctness	Delayed post-test correctness	Pre-test to post-test	Pre-test to delayed post-test
29	58,33 %	62,50 %	45,83 %	4,17 %	-12,50 %
28	54,17 %	62,50 %	66,67 %	8,33 %	12,50 %
30	54,17 %	75,00 %	66,67 %	20,83 %	12,50 %
30	50,00 %	66,67 %	41,67 %	16,67 %	-8,33 %
36	83,33 %	91,67 %	87,50 %	8,34 %	4,17 %
34	54,17 %	58,33 %	66,67 %	4,16 %	12,50 %
17	66,67 %		50,00 %		-16,67 %
	41,67 %	62,50 %	37,50 %	20,83 %	-4,17 %
28	54,17 %	54,17 %	66,67 %	0 %	12,50 %
33	87,50 %		79,17 %		-8,33 %
36	54,17 %				
33	54,17 %	50,00 %		-4,17 %	
32	66,67 %	54,17 %		-12,50 %	
17	50,00 %	62,50 %		12,50 %	
30	45,83 %				

Table 23 reveals some interesting findings. Firstly, the results show some correlation between proficiency score and accuracy percentage. Looking at the pre-test and proficiency score, one can see that only two participants had a near-target like accuracy percentage above 90 %. These participants had an accuracy percentage of 83,33 and 87,50, with proficiency scores of 33 and 36, respectively. However, there seems to be little correlation between proficiency score and accuracy percentage for the rest of the participants. For example, one participant had a 17 proficiency score, one of the two lowest scores in this group, and a 66,67 accuracy percentage, which is among the highest accuracy percentages in the pre-test for this group. Additionally, one participant had a proficiency score of 36, one of the two highest scores in the PPP group, however, with an accuracy percentage of 54,17, which is under the group's mean at 58,33 % and far from target like. Only three participants have proficiency scores above the mean (29,5) with accuracy percentages above the mean. Accordingly, there seems to be no correlation between proficiency score and accuracy in the pre-test group.

When comparing the accuracy percentage of the pre-test and post-test, there is an increase in accuracy percentage in eight participants, no improvement in one participant and a decrease in accuracy in two participants. Four participants increase their accuracy by more than 10 %. These findings may indicate a short-term improvement effect for most participants after the intervention in the PPP group. There are, however, no clear correlations between proficiency score and accuracy percentage improvement or decrease.

The results paint a different picture when comparing the delayed post-test results to the pre-test results. Five out of the ten participants experienced decreased accuracy percentage, while the other half improved from the pre-test to the delayed post-test. Two participants experienced a decrease in accuracy above 10 %, and four participants increased their accuracy by more than 10 % in the delayed post-test. The increases and decreases in accuracy among the participants in the PPP group seem arbitrary, so there appears to be no long-term effect of the intervention. Additionally, there seems to be no correlation between proficiency score and accuracy percentage in the delayed post-test.

5.3.2 TBLT group results

The following table (24) shows the corresponding data to the former table (23) for the TBLT group.

Table 24: Data from the TBLT group's proficiency test and acceptability judgement tasks

Proficiency test scores	Pre-test correctness	Post-test correctness	Delayed post-test correctness	Pre-test to post-test	Pre-test to delayed post-test
34	54,17 %	50,00 %		-4,17 %	
35	87,50 %		91,67 %		4,17 %
31	66,67 %	75,00 %	87,50 %	8,33 %	20,83 %
29	50,00 %	41,67 %	50,00 %	-8,33 %	0,00 %
30	79,17 %	54,17 %	83,33 %	-25,00 %	4,16 %
16	66,67 %	41,67 %	58,33 %	-25,00 %	-8,34 %
34	66,67 %	62,50 %	62,50 %	-4,17 %	-4,17 %
21	87,50 %	54,17 %	41,67 %	-33,33 %	-45,83 %
34	66,67 %	50,00 %	50,00 %	-16,7 %	-16,67 %
25	50,00 %	54,17 %	54,17 %	4,17 %	4,17 %
28	62,50 %		54,17 %		-8,33 %
	58,33 %	45,83 %	58,33 %	-12,50 %	0,00 %
25	66,67 %	50,00 %		-16,67 %	
36	79,17 %	75,00 %		-4,17 %	
	58,33 %				
34	50,00 %	62,50 %		12,50 %	
35	58,33 %	41,67 %		-16,66 %	
32	62,50 %				
33	54,17 %				

Table 24 shows the results from the TBLT group. Four participants had a near-target-like accuracy percentage in the pre-test in this group. These participants had accuracy percentages of 87,50, 79,17, 87,50 and 79,17 with proficiency scores at 35, 30, 21 and 36, respectively. The mean proficiency test score for this group was 30,1176, meaning that three of the participants with high accuracy percentage also had proficiency scores near or above the mean, while one of these participants had a proficiency score far below the mean. Both participants with proficiency scores below and above the mean have accuracy percentages below or above the mean (64,47 %) in the pre-test. Consequently, similarly to the PPP group, there seems to be little correlation between proficiency test score and pre-test accuracy percentage for this group.

Comparing the pre-test and post-test results reveal divergent results from the PPP group. Out of 14 participants in the post-test, only three have improved their accuracy from the pre-test, while the rest have decreased their accuracy percentage. In addition, seven participants decreased their score by over 10 %, two of these by over 20 %, and one participant decreased their accuracy by 33,33 %. These findings indicate a negative short-term effect from the intervention, as opposed to the results from the PPP group. Similarly, however, there appears to be no correlation between accuracy development from the pre-test to post-test and proficiency score to the PPP group.

Five participants had improved their accuracy percentage in the delayed post-test, and five had decreased their accuracy percentage from the pre-test. In contrast, one participant maintained the same accuracy percentage. One participant improved their accuracy score in the delayed post-test by over 20 %, while one experienced an above 10 % decrease in accuracy. The most prominent number from the delayed post-test is from one participant whose accuracy decreased by 45,83 % from the pre-test to the delayed post-test. Accordingly, there are no apparent long-term effects from the intervention, and the changes in accuracy seem arbitrary. Similarly to the PPP group results, there seems to be no correlation between proficiency score and pre-test to delayed post-test accuracy development.

The changes in accuracy across the three tests seem arbitrary for both groups, so no short- or long-term effects are evident. Furthermore, no correlation was found between proficiency score and development in accuracy percentage across tests in either group. Consequently, prediction 3 is not born out for these groups.

5.4 Test scores and intervention effects

This section examines the intervention effects based on the test results from the pre-test, the post-test and the delayed post-test. The results are discussed concerning prediction 4 and prediction 5 in 5.4.1 and 5.4.2.

- **Prediction 4**: The post-test will show improvement in test scores for both groups.

- **Prediction 5**: Improvements in the PPP group will subside in the delayed post-test, while the improvements will remain in the TBLT group. Alternatively, the improvements in the TBLT group subside less than the improvements in the PPP group.

5.4.1 Comparing the test scores

Figure 4 shows the variation in accuracy percentage throughout all acceptability judgement tasks for both groups.

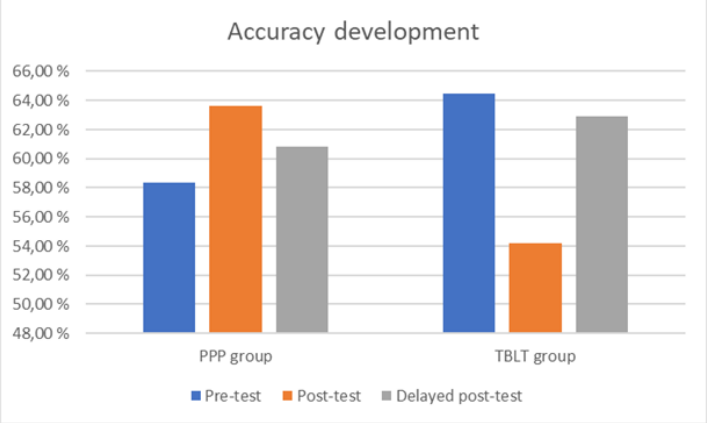


Figure 4: Accuracy development for both groups, across all three tests

Figure 4 shows that the group's accuracy percentage development over time varies greatly. The PPP group's accuracy percentage goes from 58,33 in the pre-test to 63,64 in the post-test and to 60,84 in the delayed post-test. By contrast, the TBLT group's accuracy percentage goes from 64,47 in the pre-test, 54,17 in the post-test, and 62,88 in the delayed post-test. Consequently, the PPP group increased their accuracy, short-term, by 5,31 % and long-term, by 2,51 %. Conversely, the TBLT group decreased their accuracy short-term, by 10,30 % and long-term, by 1,59 %. The following sub-sections will examine these findings concerning prediction 4 and prediction 5.

Prediction 4 stated that both groups would improve test scores in the post-test. This was only true for the PPP group and opposite of the results from the TBLT group in the post-test, where

the group experienced a decrease in accuracy percentage. As the two groups were subjected to different interventions – a PPP approach and a TBLT session, the former seems to have a more significant short-term effect in teaching subject-verb agreement (elaborated upon in section 7.2.3). The difference is striking; however, when considering the discrepancy in participation across the acceptability judgement tasks, the results have little reliability in determining group performances. This discrepancy must also be considered when analyzing the results in relation to prediction 5.

5.4.2 Long-term development

The variation between the two groups is smaller when examining the delayed post-test results. The PPP group sees a decrease in accuracy from the post-test but an increase from the pre-test. This decrease may suggest that the short-term effects of the intervention have subsided to some degree. In contrast, the TBLT group's accuracy increased from the post-test and is more akin to the group's results from the pre-test. On the other hand, the delayed post-test results from the TBLT group are still lower than the results from the pre-test, indicating that the intervention had no long-term effects. **Prediction 5** stated that the improvements made by the PPP group would subside or subside more than the improvements made by the TBLT group, in the delayed post-test. As can be seen in figure 4, this did not occur. Although the PPP group's improvements did subside from the post-test to the delayed post-test, this was only a part of the prediction. Because the TBLT-group's accuracy percentage decreased from the pre-test to the delayed post-test, **prediction 5** was ultimately contradicted. This analysis must be considered with the participation discrepancy in mind, so the results are not representative of what might have been if all participants of both groups participated in each test.

5.5 The interventions – sequence of events

In this sub-section, the sequence of events for both interventions are presented. In contrast to the descriptions of the interventions, presented in 4.3.1 and 4.3.2, this sub-section presents how the two teaching sessions unfolded, from the perspective of the instructor and the participants.

5.5.1 Presentation Practice Production

The presentation of the grammatical constructions went well, and the students seemed to pay attention at first; however, they seemed to lose attention towards the end of the presentation. After the presentation, the students reported (prompted by the instructor) that they found the presentation tedious and difficult when the constructions became more complicated.

During the practice stage, the students seemed to find some of the situational grammar tasks challenging, especially those with long-distance agreement. The students also seemed to find this task a bit boring. The next task was, reportedly, more enjoyable, and the students seemed to like the competitive aspect.

During the production stage of the session, the students were given pictures and asked to describe the actions present in the pictures, hoping to produce the targeted grammatical construction. Unfortunately, the participants produced fewer sentences containing the targeted structure than expected. However, all groups produced some sentences and managed to discuss them with each other. In sum, the session was mostly successful, and the students seemed to enjoy the assignments when they went beyond just grammar-related tasks or instruction.

5.5.2 Task Based Language Teaching

The priming activity went according to plan. First, the instructor and the students discussed the term “philanthropy”, which lasted for about five minutes. The students mostly used YouTube related celebrities as examples, and the instructor prompted discussions on the positive and negative sides of philanthropy.

The task received mixed receptions within the group. It was clear that about half of the students did not put much effort into their work. This may be described as one of the weaknesses of the intervention; the intervention is an isolated event with no actual impact on the students grading. Consequently, a substantial part of the students did not consider the task worth the effort. However, the task worked as intended for the students who did put in the effort. That means the instructor got to view the participant’s work while in progress and correct ungrammatical constructions of the targeted structure where it was needed. The post-task went as intended, and the group seemed to pay attention to the instruction.

6 Limitations

This chapter discusses sources of error and possible ways to improve the study. In 6.1, problems related to timing are discussed, while 6.2 discusses participation deficits in the study. 6.3 discusses why some of these problems occurred and ways these problems could have been avoided. Lastly, 6.4 discusses the limitations and strengths of doing the study in a naturalistic setting.

6.1 Timing

The quasi-experiment (see 4) was designed to last one month. During week one, the pre-test, intervention and the post-test were performed. The study intended to do the delayed post-test during week five. However, due to logistical problems, i.e. the lower secondary school where the study took place could not provide access to the groups as per the design, which led to the delayed post-test being performed in week eight. Having the delayed post-test three weeks after the intended time interval is not a problem. Measuring the effects on a long-term basis might even be an advantage. On the other hand, if the purpose were to measure the effects of the intervention through a larger time span, more post-tests would have been preferable, as more test-results could highlight patterns of diminishing or increasing effects.

6.2 Participants

Another problem faced during this study was participants not attending class at the times when the tests were performed. This led to far fewer participants having their knowledge measured in all tests. As stated earlier, only 8 out of 16 participants were measured at all points in the PPP group, and only 11 out of the 20 participants in the TBLT group were measured at all points. In other words, far fewer data was gathered than what was intended. Because of this, the representability of the data collected during the study is limited. To improve the representability of the study, a larger group should be used. In this case, that means using several classes, perhaps across several schools.

6.3 Why these problems occur and possible countermeasures

One of the risks of doing a study in one or more school classes is the unpredictability that follows. Regarding timing, the long delay in performing the delayed post-test according to plan was because of spontaneous issues that could not be pre-planned for. Having the students follow their curriculum is a priority, and if the agreed-upon appointments for test-taking stand in the way of this, motivation for study participation might be impaired. Low student participation

had several causes. Sometimes, a participant had other duties to attend to at the appointed hour. Other times; there was sickness. There were also times when some participants were unmotivated and refused to take the test.

These are complex issues to prevent, as they are a natural part of everyday school. There are, however, measurements that could have been taken to limit participant dropout. Suppose the study had been organized as part of a more substantial cooperative venture. In that case, the school from which the participant data was collected might have been more inclined to take measures to facilitate a more fluid run of the study. In other words, closer cooperation between schools and universities might lay the groundwork for smoother operations.

6.4 Naturalistic setting

This study was conducted in a naturalistic setting, and therefore, it is affected by several uncontrolled variables (see 4). The groups were chosen based on the classes the participants were in, i.e., intact classes, so the instruction they received throughout their English education is diverse. Subsequently, this may factor into the two group's English proficiency levels. There will also be differences in learning and social environment between the two groups. On the other hand, these factors are always a part of Norwegian's formal education. Therefore, the study reflects the school setting realistically, increasing the face validity of the study (see 4).

7 Discussion

This section discusses the research questions and predictions considering the study's results and the background presented in section 2. The research questions and predictions were the following:

> *RQ 1: Do Norwegian L1 learners of English have little knowledge of the third person singular "-s"?*
>
> *RQ 2: Is there any variation in accuracy on the sub-categories; local singular agreement, local plural agreement, long-distance singular agreement and long-distance plural agreement among L1 Norwegian L2 English learners?*
>
> *RQ 3: Does instruction affect learner's knowledge of subject verb agreement?*
>
> *RQ 4: Do different types of instruction affect learner's knowledge of subject-verb agreement differently?*

Prediction 1: Subject-verb agreement will be problematic for both groups in the pre-test.

Prediction 2: The sub-conditions local plural agreement and long-distance singular agreement will be more challenging for the participants than local singular agreement. Long-distance plural agreement will be the most difficult.

Prediction 3: Participants with low proficiency test scores will improve more in the post-test and delayed post-test than those with high proficiency test scores.

Prediction 4: The post-test will show improvement in test scores for both groups.

Prediction 5: Improvements in the PPP group will subside in the delayed post-test, while the improvements will remain in the TBLT group. Alternatively, the improvements in the TBLT group subside less than the improvements in the PPP group.

In 7.1, the predictions are reviewed in a discussion of the results in section 5. The research questions are addressed in 7.2, connecting this study's findings to the framework presented in the background section (2). Lastly, 7.3 discusses the possible pedagogical implications of the study.

7.1 The predictions

7.1.1 Subject-verb agreement will be problematic for both groups in the pre-test.

Prediction 1 is corroborated because both groups had a lower than 90 % accuracy (see 5.4.1), which is considered target-like. Individually, only two participants scored above 90 % accuracy, one from the PPP group in the post-test (see table 23) and one from the TBLT group in the delayed post-test (see table 24). Additionally, none of the groups scored higher than 90 % in any of the sub-conditions throughout all three tests. This is in line with the previously discussed studies: Jensen (2016), Jensen et al. (2019) and Garshol (2018). Some participants achieved near-target like scores throughout the tests, however, most participants scored considerably lower.

7.1.2 Plural and Long-distance agreement is more challenging than local singular agreement.

The challenge posed by plural and long-distance agreement was formulated in **prediction 2** and follows findings in Jensen et al. (2019). The participants in this study found long-distance plural agreement to be the most challenging sub-condition. This trend remained constant throughout all three tests. On the other hand, the data was not as concise about the remaining sub-conditions. Local singular agreement usually had the highest accuracy in both groups, however, not consistently. For example, in the pre-test, the TBLT-group had higher accuracy in the long-distance singular sub-condition than in the local singular sub-condition. Hence, **prediction 2** does not bear out completely.

7.1.3 Low proficiency and test score improvement

Participants with lower proficiency scores were predicted to improve more after the intervention in contrast to participants with higher proficiency scores (prediction 3). This tendency was predicted based on earlier research (Jensen, 2016), where there was found to be a correlation between low proficiency scores and low accuracy in acceptability judgement. This correlation might indicate that participant's with low proficiency have little knowledge of agreement and might have the most to gain from instruction.

The changes in accuracy across the three tests seem arbitrary for both groups, so no short- or long-term effects are evident. Furthermore, no correlation was found between proficiency score

and development in accuracy percentage across tests in either group. Consequently**, prediction 3** is not born out for these groups.

7.1.4 Short-term improvements in accuracy

Prediction 4 stated that both groups would see improvements from the pre-test to the post-test. The results, however, show that the TBLT group did not experience any improvements in the post-test. On the contrary, the TBLT group's accuracy decreased shortly after the intervention. Then again, the PPP group's results did increases substantially in the post-test, indicating a short-term effect from the corresponding intervention (see 5.4.1; figure 4). Accordingly, **prediction 4** is somewhat fulfilled as short-term effects from the intervention can be observed in the PPP group.

7.1.5 Long-term improvements in accuracy

Prediction 5 professed that improvements observed in the post-test would remain for the TBLT group and subside for the PPP group, or alternatively, that improvements would subside less for the TBLT group. Either way, the prediction implied that the TBLT intervention would have more robust long-term effects than the PPP intervention. The data revealed that there were no significant long-term effects for either group (see 5.4.1; figure 4). The TBLT group did increase their accuracy to near pre-test results, back up from the groups substantial decrease in the post-test, however, the intervention is no apparent causal factor. Furthermore, the increase in accuracy observed in the post-test fell back to near-pre-test levels for the PPP group, thus the effects from the intervention do not seem to carry on long-term. Consequently, **prediction 5** was not corroborated. Note however, that the PPP group, which initially had a lower overall accuracy in the pre-test, ended up catching up to the TBLT group in the delayed post-test.

7.2 The research questions

This sub-section seeks to answer the abovementioned research questions by, firstly, discussing RQ 1 and RQ 2 through the findings from section 5 with empirical data and theory from the GenSLA research field, presented in section 2. The later research questions (RQ 3 and RQ 4) are discussed with theory on ISLA as its backdrop, also presented in section 2.

7.2.1 Do Norwegians have little knowledge of the third person singular "-s"?

This study adopts the views presented in White (2003) and Slabakova (2008), known as "syntax-before-morphology" and "the bottleneck hypothesis", respectively. Both see functional

morphology as more challenging to acquire than syntax and view functional morphology a stumbling block for English learners at more advanced levels. These views are supported in the data from Jackson et al. (2018) for Chinese learners, making the authors conclude that functional morphology is challenging regardless of the learner's L1. The present study sought to test if L1 Norwegian learners of L2 English in the lower secondary school followed this pattern by testing their accuracy in judging sentences with subject-verb agreement.

The accuracy scores found among the participants from this study are echoed in Jensen et al. (2019.) This study set out to test the bottleneck hypothesis and found that the third person singular "-s" seems harder to acquire than non-V2 syntax, which supports the hypothesis. Similar evidence was also found in Jensen (2016), however, only with respect to judging ungrammatical sentences. Garshol (2018) did not seek to test the bottleneck hypothesis, however, the study did find evidence from corpus data supporting the claim that L1 Norwegians find English agreement challenging. The bottleneck hypothesis was also adopted in this thesis, and accordingly, participants were predicted to find subject-verb agreement challenging, however, not by comparing it to verb movement like in Jensen et al. (2019).

Throughout the present study, both groups scored, on average, substantially lower than what is described as target like, and there were only a few individual cases of target like accuracy. There was, however, no correlation between general proficiency level and acceptability judgement task score. Therefore, the participant's acquisition of the third person singular "-s" cannot be categorized as relatively low in relation to their general English proficiency level. Nevertheless, the group's accuracy scores are far from target like, making it likely that the participants have relatively limited knowledge of the third person singular "-s". Consequently, this study maintains the bottleneck hypothesis claiming the participants did indeed find subject-verb agreement challenging.

7.2.2 Variations in accuracy on the sub-conditions

In Jensen et al. (2019) the participants showed lower accuracy in some sub-conditions than others. As mentioned above, this was also predicted to occur in the present study. Looking at the data from 5.2, it is apparent that the long-distance plural sub-condition was the most challenging for the participants in this study. The other sub-conditions were less conclusive; however, local singular agreement was usually the easiest of the four sub-conditions. The cause for this tendency in Norwegian learners is explored in the following.

This discrepancy between accurately judging the different sub-conditions shows that subject-verb agreement, a challenging concept to learn in the first place, can become even more difficult when presented in sentences with more "noise". That means, when the subject and the verb are separated by longer distances, it seems that the participants find it harder to judge the sentence accurately. The long-distance plural sub-condition has been recorded in earlier research as the most difficult of the sub-conditions, and so plurality also appear to play a role. One reason as to why plurality may impact accuracy judgement is because plural agreement has no agreement marker, and so the singular, "-s"-marking, is overgeneralized. Jensen et al. (2019, 21) hypothesize that this overgeneralization may be due to overlearning of the singular "-s"-marker. Furthermore, the authors also reason that processing difficulties make the long-distance plural sub-condition challenging, because it provides two elements of difficulty into the structure (22), as opposed to local plural and long-distance singular agreement, which only have one element of difficulty.

7.2.3 The effects of language instruction

This study used an intervention on two different groups to test the effects of instruction on the learner's knowledge of subject-verb agreement. The study was used to answer RQ 3 and RQ 4. The former posed the question of whether instruction influences subject-verb agreement in L1 Norwegian L2 English learners, and the latter asked whether different instructional methods have differing effects. The PPP group's intervention was a teaching method called TBLT (see 2.3.7). One of the principles of grammar teaching through tasks is to teach it implicitly. PPP, on the other hand (see 2.3.8), teaches grammar explicitly. To discuss the research questions, the results from the present study, empirical evidence and literature from the ISLA field of research is used.

7.2.3.1 Does instruction affect learner's knowledge of subject-verb agreement?

The results in the present study display varied effects for each group (see 5.4.1). The post-test results show a short-term increase in accuracy in the PPP group, while the TBLT group experienced a substantial decrease. Conversely, in the delayed post-test, both groups exhibit similar results to the ones found in the pre-test, showing no convincing long-term effects from the intervention. This means that for these two groups, the instruction did not affect the learner's knowledge of subject-verb agreement in the long-term. It should, however, be noted that the PPP-group did catch up to the TBLT group.

Based on these findings, the two groups responded somewhat differently to the two interventions used. In Hirakawa, Shibuya & Endo (2019) (see 2.3.2) two groups of Japanese students were either exposed to explicit input or naturalistic input. Here, the group that received explicit instruction did show improvement on the targeted structure, first in the post-test, then, further development in the delayed post-test. The group receiving natural exposure did not show significant improvement in the post-test or delayed post-test. This result echoes the results from the post-test in the present study (see 5.4.1), where the group receiving explicit grammar instruction was the only group showing improvement. On the other hand, while natural exposure and TBLT both provide implicit grammar input, TBLT can still focus its tasks on specific grammatical constructions in a way that is impossible through natural exposure alone. Therefore, the comparison should be taken with a grain of salt. Nevertheless, teaching grammar through PPP, and explicit grammar instruction are comparable, and so the short-term effects that were measured in the PPP group may possibly support the findings in Hirakawa, Shibuya & Endo (2019).

In contrast, the findings from Umeda et al. (2019) showed, through multiple post-tests that long-term effects of explicit grammar instruction decreased after the instruction subsided. The improvements made by the participants subsided between post-test 2 (one week after the intervention) and post-test 3 (twelve weeks after the intervention). These results are more like those found in the present study (see 5.4.1), were the effects of the PPP-intervention subsided, between the post-test (one week after the intervention) and the delayed post-test (two months). However, the scores were still higher than in the pre-test (see figure 4, 5.4.2). This may be interpreted in two ways: Either the participants have retained some of their improved knowledge of subject-verb agreement or, the effects will continue to subside. Ideally, the study would have used more post-tests over a longer time-span to better examine the long-term effects of the intervention.

It should also be noted that the metanalysis (Li, 2010), which focused on corrective feedback (see 2.3.4), found that explicit CF was more effective in the short-term, while implicit CF was more effective in the long-term. Again, these results are reflected in the post-test results from the PPP group in the present study, and the results from the post-tests in the abovementioned studies.

Ultimately, the literature seems to favor both implicit and explicit grammar teaching along the lines of the focus on form view, however, with an emphasis on implicit learning (see 2.3.2).

Hirakawa, Shibuya & Endo (2019) and Umeda et al. (2019) come to diverging conclusions regarding the effects of explicit grammar instruction. On the other hand, they did not test the same construct, testing for adjective ordering restrictions and article use, respectively. The same can be said for using the present study in this discussion, which should be kept in mind.

7.3 Pedagogical implications

The results show that grammatical instruction might be effective in the short-term for learning subject-verb agreement; however, The PPP group was the only group that showed improvement. Accordingly, the finding may support explicit grammar teaching in English instruction. However, as mentioned in 5.5, the group subjected to the PPP intervention found the situational grammar task and the presentation explaining subject-verb agreement tedious. The students also found the presentation hard to follow. This may be because the participants found it tedious or because the participants had little explicit knowledge of subject-verb agreement; thus, the more complicated structures became a challenge for them. Nevertheless, explicit grammar teaching might be an effective way of teaching subject-verb agreement, however, instructors should be careful to design the tasks in ways that are engaging for the students. Creating engaging tasks may increase student motivation and involvement in the teaching session and, consequently, the effectiveness of instruction.

The effects of the intervention effects on the TBLT group might have been more substantial if the teaching session had been more successful. In addition, during the intervention, the main task engaged fewer students than was intended. If the study was carried out as a part of the student's curriculum and the main task was incorporated as part of their assessment basis, the task might have been more beneficial. Accordingly, the findings from the present study should be read with these shortcomings in mind.

The results from the sub-conditions showed that local singular agreement was the least challenging for the participants and that the plural and long-distance conditions made the sentences more demanding. In Jensen et al. (2019, 18), the results show that the local singular agreement sub-condition develops in a steeper slope than the other sub-conditions as proficiency scores get higher. Unsurprisingly, the long-distance plural sub-condition develops in a gentler slope, showing that a higher proficiency score was less impactful for this sub-condition. Therefore, language instruction might gain from more specific targeting of subject-verb agreement constructions with plural nouns and long-distance agreement, as even proficient learners have difficulties with these sentences.

The present study used acceptability judgement testing to compare two different pedagogical interventions, a novel design that should be used in the future. Although the findings in the present study may be lacking in representativity, its design might be improved upon to make up for the shortcoming of the present study. Using a greater sample, with more groups and larger group sizes, a similar design might make discoveries that can be used in language instruction. The design can also be improved by employing more post-tests and increasing the duration of the study to capture the long-term effects of the interventions.

8 Conclusion

The study in this thesis tested the acquisition of L2 English subject-verb agreement in two groups of L1 Norwegian lower secondary school students. Previous studies (Jensen et al., 2019; Garshol 2018) have shown that English subject-verb agreement is challenging for Norwegians. The findings from the former study also showed that plurality and long-distance make the subject-verb agreement harder to acquire for L1 Norwegians. However, few studies have investigated the effects of instruction on L1 Norwegian L2 English subject-verb acquisition. Therefore, the present study also tried to compare two pedagogical approaches to grammar instruction; Task Based Language Teaching and Presentation Practice Production.

To achieve these goals, the present study used an acceptability judgement task in a pre-test, post-test and a delayed post-test. The intervention was carried out right after the pre-test and one day before the post-test, and was made up of two different, 60-minute teaching sessions – one for each group. Additionally, an English proficiency test and a questionnaire were used to examine whether proficiency score is linked to accuracy in the acceptability judgement tasks. The questionnaire was used mainly to determine the participants L1's, gender and year of birth.

The findings from the abovementioned studies make up the basis for two of the research questions (RQ 1 and RQ 2), developed for this thesis. RQ 3 and RQ 4 were formulated based on literature from the ISLA framework (see 2.3). Thus, the following research questions were conceived: RQ 1: Do Norwegian L1 learners of English have little knowledge of the third person singular "-s"? RQ 2: Is there any variation in accuracy on the sub-categories; local singular agreement, local plural agreement, long-distance singular agreement and long-distance plural agreement among L1 Norwegian L2 English learners? RQ 3: Does instruction affect learner's knowledge of the third person singular "-s"? RQ 4: Do different types of instruction affect learner's knowledge of the third person singular "-s" differently?

The research questions were investigated through the acceptability judgement tasks. The results from the acceptability judgement varied substantially from pre-test to post-test and from post-test to delayed post-test. The test scores show that the participants do indeed find subject-verb agreement challenging, all scoring below 90 % accuracy. This is in line with the previously described studies Jensen et al. (2019) and Garshol (2018). Jensen et al. (2019) also found that plural and long-distance agreement was more challenging than local singular agreement. This

is supported by the results in the present study, where the participants had the lowest accuracy when judging the condition that targeted long-distance plural agreement.

Using the pre-test scores as a baseline, the TBLT group decreased in mean accuracy scores substantially in the post-test and slightly in the delayed post-test. The PPP group increased their accuracy substantially in the post-test and slightly in the delayed post-test. Accordingly, the PPP intervention may have had positive short-term effects on subject-verb agreement accuracy. There were no substantial long-term effects for either group. However, the results should be taken with a grain of salt because of the variation in participation for each test; additionally, the groups were initially small, so the representability of the study is limited.

Subject-verb agreement is challenging for L1 Norwegians learning L2 English. Therefore, future investigations should explore effective ways of teaching this phenomenon. Combining acceptability judgement tasks and pedagogical interventions to measure the effects of language instruction is a promising design. Future research should use larger sample sizes and use longer studies, with additional post-tests. These improvements would increase the representability and better capture the long-term effects of the interventions. The present thesis sought to highlight challenges in acquiring subject-verb agreement and, additionally, to compare implicit and explicit grammar instruction. The results were not completely in line with the predictions. However, through the study's design, these goals were ultimately reached.

Work cited

Brown, R. (1973). A First Language: The Early Stages. Cambridge: Harvard University Press.

Criado, R. (2013). A critical review of the presentation-practice-production model (PPP) in foreign language teaching. Homenaje a francisco gutiérrez díez, 97-115.

Dabrowska. (2010). Naive v. expert intuitions: An empirical study of acceptability judgments. Linguistic Review, 27(1), 1–23. https://doi.org/10.1515/tlir.2010.001

Dalili, Mehdi Vaez. "On the Integration of Form and Meaning in English Language Teaching (ELT): An Overview of Current Pedagogical Options." Procedia, Social and Behavioral Sciences, vol. 15, 2011, pp. 2117–2121

East, M. (2017). Research into practice: The task-based approach to instructed second language acquisition. Language Teaching, 50(3), 412–424. https://doi.org/10.1017/S026144481700009X

Ellis, R (2003). Task-based language learning and teaching. New York, NY: Oxford University Press

Ellis, R. (2006). Current Issues in the Teaching of Grammar: An SLA Perspective. TESOL Quarterly, 40(1), 83–107. https://doi.org/10.2307/40264512

Ellis. (2009). Task-based language teaching: sorting out the misunderstandings. International Journal of Applied Linguistics, 19(3), 221–246. https://doi.org/10.1111/j.1473-4192.2009.00231.x

Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C. (2019). Task-Based Language Teaching. Cambridge University Press. https://doi.org/10.1017/9781108643689

Garshol, Lenka. I Just Doesn't Know: Agreement Errors in English Texts by Norwegian L2 Learners: Causes and Remedies, 2019.

Hirakawa, Shibuya, M., & Endo, M. (2019). Explicit instruction, input flood or study abroad: Which helps Japanese learners of English acquire adjective ordering? Language Teaching Research: LTR, 23(2), 158–178. https://doi.org/10.1177/1362168817752377

Hummel, K. M. (2014). Introducing Second Language Acquisition Perspectives and Practices

Jackson, Mormer, E., & Brehm, L. (2018). THE PRODUCTION OF SUBJECT-VERB AGREEMENT AMONG SWEDISH AND CHINESE SECOND LANGUAGE SPEAKERS OF ENGLISH. Studies in Second Language Acquisition, 40(4), 907–921. https://doi.org/10.1017/S0272263118000025

Jensen. (2016). The Bottleneck Hypothesis in L2 acquisition: Norwegian L1 speakers' knowledge of syntax and morphology in English L2. UiT Norges arktiske universitet.

Jensen, I. N., Slabakova, R., Westergaard, M., & Lundquist, B. (2019). The Bottleneck Hypothesis in L2 acquisition: L1 Norwegian learners' knowledge of syntax and morphology in L2 English. https://doi.org/10.1177/0267658318825067

Jensen, 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, 36(1), 3–29. https://doi.org/10.1177/0267658318825067

Lardiere. (1998a). Case and tense in the 'fossilised' steady state. Second Language Research, 14(1), 1–26. https://doi.org/10.1191/026765898674105303

Lardiere. (1998b). Dissociating syntax from morphology in a divergent L2 end-state grammar. Second Language Research, 14(4), 359–375. https://doi.org/10.1191/026765898672500216

Lipner, M., Armon-Lotem, S., Walters, J., & Altman, C. (2021). Crosslinguistic Influence (CLI) of Lexical Breadth and Depth in the Vocabulary of Bilingual Kindergarten Children – A Bilingual Intervention Study. Frontiers in Psychology, 12, 671928–671928. https://doi.org/10.3389/fpsyg.2021.671928

Li. (2010). The Effectiveness of Corrective Feedback in SLA: A Meta-Analysis. Language Learning, 60(2), 309–365. https://doi.org/10.1111/j.1467-9922.2010.00561.x

Lorenz, E., Bonnie, R. J., Feindt, K., Rahbari, S., & Siemund, P. (2019). Crosslinguistic influence in unbalanced bilingual heritage speakers on subsequent language acquisition: Evidence from pronominal object placement in ditransitive clauses. International Journal of Bilingualism, 23(6), 1410–1430. https://doi.org/10.1177/1367006918791296

Loewen, S. & Philp, J. (2012), Instructed Second Language Acquisition. In Mackey, A., & Gass, S. M. (Eds)., (2012). Research Methods in Second Language Acquisition (1. Aufl.). (pp. 53-73). Wiley-Blackwell.

Malcolm, T. R. (2021). Cross-Linguistic Morphosyntactic Influence InBilingual Speakers of Jamaican Creole and Jamaican English. ProQuest Dissertations Publishing.

Mackey, & Gass, S. M. (2005). Second language research: methodology and design (pp. XV, 405). Lawrence Erlbaum.

Rodrigo, Krashen, S., & Gribbons, B. (2004). The effectiveness of two comprehensible-input approaches to foreign language instruction at the intermediate level. System (Linköping), 32(1), 53–60. https://doi.org/10.1016/j.system.2003.08.003

Rothman, J., & Slabakova, R. (2018). THE GENERATIVE APPROACH TO SLA AND ITS PLACE IN MODERN SECOND LANGUAGE STUDIES. Studies in Second Language Acquisition, 40(2), 417–442. https://doi.org/10.1017/S0272263117000134

Schwartz, & Sprouse, R. A. (1996). L2 cognitive states and the Full Transfer/Full Access model. Second Language Research, 12(1), 40–72. https://doi.org/10.1177/026765839601200103

Serratrice, L. (2013). Crosslinguistic influence in bilingual development Determinants and mechanisms. Linguistic Approaches to Bilingualism, 3(1), 3–25. https://doi.org/10.1075/lab.3.1.01ser

Shintani. (2016). Input-based tasks in foreign language instruction for young learners (Vol. 9). John Benjamins Publishing Company.

Slabakova, R (2008) Meaning in the second language. Berlin: Mouton de Gruyter.

Slabakova, Roumyana (2016). Second Language Acquistion. Oxford University Press, Oxford.

Slabakova, R. (2017). The scalpel model of third language acquisition. The International Journal of Bilingualism: Cross-Disciplinary, Cross-Linguistic Studies of Language Behavior, 21(6), 651–665. https://doi.org/10.1177/1367006916655413

Slabakova. (2019). The Bottleneck Hypothesis Updated. In Three Streams of Generative Language Acquisition Research. John Benjamins Publishing.

Smith, & Truscott, J. (2014). Explaining input enhancement: a MOGUL perspective. International Review of Applied Linguistics in Language Teaching, IRAL, 52(3), 253–281. https://doi.org/10.1515/iral-2014-0012

Spada, Nina. "SLA Research and L2 Pedagogy: Misapplications and Questions of Relevance." Language Teaching, vol. 48, no. 1, 2015, pp. 69–81.

Toth, Paul D, and Moranski, Kara. "Why Haven't We Solved Instructed SLA? A Sociocognitive Account." Foreign Language Annals, vol. 51, no. 1, 2018, pp. 73–89

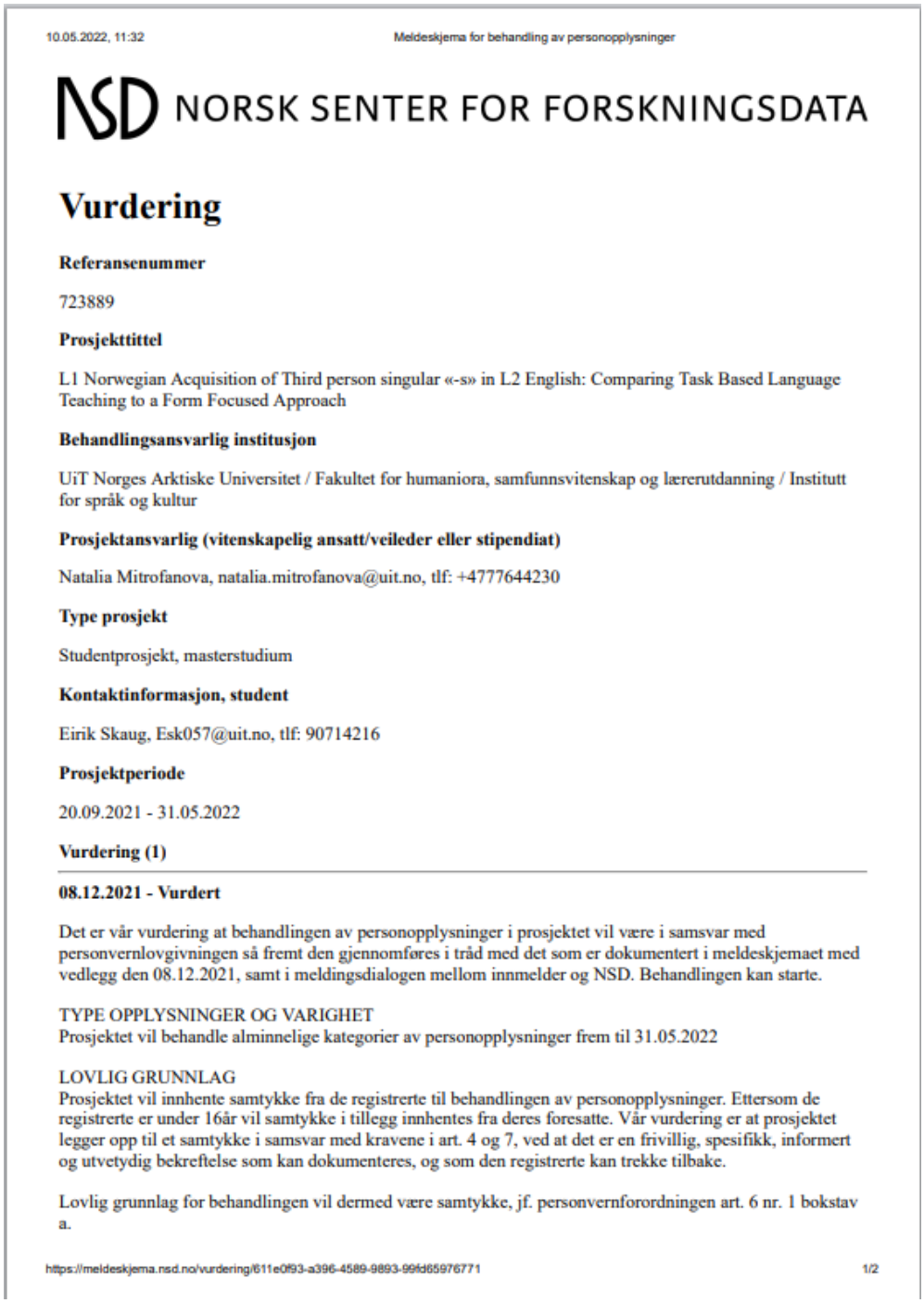
Umeda, Snape, N., Yusa, N., & Wiltshier, J. (2019). The long-term effect of explicit instruction on learners' knowledge on English articles. Language Teaching Research: LTR, 23(2), 179–199. https://doi.org/10.1177/1362168817739648

Westergaard, M., Mitrofanova, N., Mykhaylyk, R., & Rodina, Y. (2017). Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. The International Journal of Bilingualism: Cross-Disciplinary, Cross-Linguistic Studies of Language Behavior, 21(6), 666–682. https://doi.org/10.1177/1367006916648859

White, L. (2003). Morphological variability and the morphology/syntax interface. In Second Language Acquisition and Universal Grammar (Cambridge Textbooks in Linguistics, pp. 178-202). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511815065.007

Wang. (2019). The influence of creative task engagement on English L2 learners' negotiation of meaning in oral communication tasks. System (Linköping), 80, 83–94. https://doi.org/10.1016/j.system.2018.10.015

Appendix 1: Approval from NSD to conduct the study



NSD NORSK SENTER FOR FORSKNINGSDATA

Vurdering

Referansenummer

723889

Prosjekttittel

L1 Norwegian Acquisition of Third person singular «-s» in L2 English: Comparing Task Based Language Teaching to a Form Focused Approach

Behandlingsansvarlig institusjon

UiT Norges Arktiske Universitet / Fakultet for humaniora, samfunnsvitenskap og lærerutdanning / Institutt for språk og kultur

Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)

Natalia Mitrofanova, natalia.mitrofanova@uit.no, tlf: +4777644230

Type prosjekt

Studentprosjekt, masterstudium

Kontaktinformasjon, student

Eirik Skaug, Esk057@uit.no, tlf: 90714216

Prosjektperiode

20.09.2021 - 31.05.2022

Vurdering (1)

08.12.2021 - Vurdert

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 08.12.2021, samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan starte.

TYPE OPPLYSNINGER OG VARIGHET
Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 31.05.2022

LOVLIG GRUNNLAG
Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Ettersom de registrerte er under 16år vil samtykke i tillegg innhentes fra deres foresatte. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte kan trekke tilbake.

Lovlig grunnlag for behandlingen vil dermed være samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a.

PERSONVERNPRINSIPPER
NSD vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om:

· lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte og de foresatte får tilfredsstillende informasjon om og samtykker til behandlingen
· formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke behandles til nye, uforenlige formål
· dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet
· lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER
Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), og dataportabilitet (art. 20).

NSD vurderer at informasjonen om behandlingen som de registrerte og de foresatte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.

Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER
NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32).

Ved bruk av databehandler (spørreskjemaleverandør, skylagring eller videosamtale) må behandlingen oppfylle kravene til bruk av databehandler, jf. art 28 og 29. Bruk leverandører som din institusjon har avtale med.

For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og/eller rådføre dere med behandlingsansvarlig institusjon.

MELD VESENTLIGE ENDRINGER
Dersom det skjer vesentlige endringer i behandlingen av personopplysninger, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. Før du melder inn en endring, oppfordrer vi deg til å lese om hvilke type endringer det er nødvendig å melde: https://www.nsd.no/personverntjenester/fylle-ut-meldeskjema-for-personopplysninger/melde-endringer-i-meldeskjema
Du må vente på svar fra NSD før endringen gjennomføres.

OPPFØLGING AV PROSJEKTET
NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.
Lykke til med prosjektet!

Appendix 2: Consent form

Vil du delta i forskningsprosjektet

"L1 Norwegian Acquisition of Third person singular «-s» in L2 English: Comparing Task Based Language Teaching to a Form Focused Approach "?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å undersøke tilegnelse av engelsk språk. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål
Formålet med denne undersøkelsen er å utforske norske elevers tilegnelse av engelsk tredjeperson «*-s*» ved å sammenligne to ulike læringstilnærminger (Task based language teaching og Presentation, practice, production).
I denne studien skal utvalgets presisjon når det gjelder å bedømme korrekte eller feilaktige setninger undersøkes. Dette gjøres for å besvare følgende spørsmål: «1) Har norske elever lav presisjon i sammen heng med å bedømme korrekt eller feilaktig bruk av samsvar, relativt til deres generelle kunnskaper i engelsk? 2) Er det en systematisk, kortsiktig eller langsiktig, endring i utvalgets skåre etter intervensjonen? Og, 3) Er det en systematisk forskjell i skåre mellom gruppene utsatt for ulike læringstilnærminger?
Dette er datainnsamling til en mastergradsoppgave i Engelsk språk.

Hvem er ansvarlig for forskningsprosjektet?
[Universitetet i Tromsø Norges Arktiske Universitet] er ansvarlig for prosjektet.

Hvorfor får du spørsmål om å delta?
Utvalget er valgt ut ifra undertegnedes praksisskole og praksisveileder, bestemt av Universitetet i Tromsø.

.

Hva innebærer det for deg å delta?

Å delta i denne undersøkelsen går ut på å anonymt besvare: 1) Engelsk språkferdighetstest, og 2) Test i grammatisk korrekthet (Hvilke setninger er akseptable og ikke) – 4 ganger i løpet av ca. 1,5 – 2 måneder.

Hvis du velger å delta i prosjektet, innebærer det at du fyller ut et spørreskjema. Det vil ta deg ca. 5 minutter. Spørreskjemaet inneholder spørsmål om: 1) hvor gammel du er, og 2) Dine kunnskaper om andre språk enn norsk og engelsk, f.eks. om du har et annet morsmål. Dine svar fra spørreskjemaet blir registrert elektronisk

Foreldre kan se spørreskjema og tester på forhånd ved å ta kontakt med prosjektansvarlig

Det er frivillig å delta
Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykket

tilbake uten å oppgi noen grunn. Alle dine personopplysninger vil da bli slettet. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

Forsøket vil ikke ha noen innvirkning på karakterer eller annet som har med skolen å gjøre.

De som ikke ønsker å være med på undersøkelsen vil få alternativt opplegg

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger
Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket.

- De som vil ha tilgang til dataen som blir samlet inn via undersøkelsene er: Undertegnede, to veiledere ved Universitetet i Tromsø.
- Navnet og kontaktopplysningene dine vil jeg erstatte med en kode som lagres på egen navneliste adskilt fra øvrige data.

Deltakerne vil ikke kunne gjenkjennes i publikasjonen av dataen, da denne ikke inneholder beskrivende informasjon, annet enn statistikk av gruppens test-resultater – som er anonym.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?
Opplysningene anonymiseres når prosjektet avsluttes/oppgaven er godkjent, noe som etter planen er: Mai, 2022. Ellers slettes personopplysningene ved prosjektslutt.

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

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

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

På oppdrag fra Universitetet i Tromsø Norges Arktiske Universitet har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

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

- Universitetet i Tromsø ved Eirik Skaug, esk057@uit.no, tlf: 907 14 216
- Veileder: Christopher Loe Olsen, christopher.l.olsen@uit.no
- Veileder: Natalia Mitrofanova, natalia.mitrofanova@uit.no
- Vårt personvernombud: Joakim Bakkevold, personvernombud@uit.no

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

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

Med vennlig hilsen

Christopher Loe Olsen
(Forsker/veileder)

Eirik Skaug

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet «L1 Norwegian Acquisition of Third person singular «-s» in L2 English: Comparing Task Based Language Teaching to a Form Focused Approach», og har fått anledning til å stille spørsmål. Jeg samtykker til:

- ☐ å delta i «acceptability judgement test"
- ☐ å delta i «proficiency test»
- ☐ å delta i spørreskjema om alder og språkkunnskaper

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

---

(Signert av prosjektdeltaker, dato)

Appendix 3: Oxford Proficiency Test

Instructions: Please complete the sentences by selecting the best answer from theavailable answers below. You can select by underlining or making an X next to your choice.

1) Water _____ at a temperature of 100° C.

is to boil | is boiling | boils

2) In some countries _____ very hot all the time.

there is | is | it is

3) In cold countries people wear thick clothes _____ warm.

for keeping | to keep | for to keep

4) In England people are always talking about _____.

a weather | the weather | weather

5) In some places _____ almost every day.

it rains | there rains | it raining

6) In deserts there isn't _____ grass.

the | some | any

7) Places near the Equator have _____ weather even in the cold season.

a warm | the warm | warm

8) In England _____ time of year is usually from December to February.

coldest | the coldest | colder

9) _____ people don't know what it's like in other countries.

The most | most of | most

10) Very _____ people can travel abroad.

less little few

11) Mohammed Ali _____ his first world title fight in 1960.

has won won is winning

12) After he _____ an Olympic gold medal, he became a professional boxer.

had won have won was winning

13) His religious beliefs _____ change his name when he became a champion.

have made him made him to made him

14) If he _____ lost his first fight with Sonny Liston, no one would have been surprised.

has would have had

15) He has traveled a lot _____ as a boxer and as a world-famous personality.

both and or

16) He is very well known _____ the world.

all in all over in all

17) Many people _____ he was the greatest boxer of all time.

is believing are believing believe

18) To be the best _____ the world is not easy.

from in of

19) Like any top sportsman, Ali _____ train very hard.

had to must should

20) Even though he has now lost his title, people _____ always remember him as a champion.

would will did

Part II:

Read the following passage about the history of aviation and choose the best answer for each blank. Note that it is a continuous story.

21) The history of _____ is

airplane the airplane an airplane

22) _____ short one. For many centuries men

quite a a quite quite

23) _____ to fly, but with

are trying try had tried

24) _____ success. In the 19th century a few people

little few a little

25) succeeded _____ in balloons. But it wasn't until

to fly in flying into flying

26) the beginning of _____ century that anybody

last next that

27) _____ able to fly in a machine

were is was

28) _____ was heavier than air, in other words, in

who which what

29) _____ we now call a 'plane'. The first people to achieve

who which what

30) 'powered flight' were the Wright brothers. _____ was the machine

his their theirs

31) which was the forerunner of the Jumbo jets and supersonic airliners that are

_____ common

such such a some

32) sight today. They _____ hardly have imagined that in 1969,

could should couldn't

33) _____ more than half a century later,

not much not many no much

34) a man _____ landed on the moon.

will be had been would have

35) Already _____ is taking the first steps towards the stars.

a man man the man

36) Although space satellites have existed _____ less

since during for

37) than forty years, we are now dependent _____ them for all

from of on

38) kinds of _____. Not only

informations information an information

39) _____ being used for scientific research in

are they they are there are

40) space, but also to see what kind of weather _____.

is coming comes come

Appendix 4: Score Sheet

List 1

Navn: _____

	GOOD	BAD
1.		
2.		
3.		
4.		
5.		
6.		
7.		
8.		
9.		
10.		
11.		
12.		
13.		
14.		
15.		
16.		
17.		
18.		
19.		
20.		
21.		
22.		
23.		
24.		

	GOOD	BAD
25.		
26.		
27.		
28.		
29.		
30.		
31.		
32.		
33.		
34.		
35.		
36.		
37.		
38.		
39.		
40.		
41.		
42.		
43.		
44.		
45.		
46.		
47.		
48.		

1

Appendix 5: Agreement sentence pairs

1) Main clauses with local agreement, singular subjects

The girl drinks a lot of water every day	*The girl drink a lot of water every day
The boy likes to go swimming in the ocean	*The boy like to go swimming in the ocean
The girl drives to work every Wednesday morning	*The girl drive to work every Wednesday morning
The student loves to read books about football	*The student love to read books about football
The teacher eats fish for dinner every Friday	*The teacher eat fish for dinner every Friday
The brown dog plays with the yellow football	*The brown dog play with the yellow football

2) Main clauses with local agreement, plural subjects

The kids like to play in the park every weekend	*The kids likes to play in the park every weekend
The teachers give their students a lot of homework	The teachers gives their students a lot homework
The cats play with the yellow and green ball	*The cats plays with the yellow and green ball
The students sit in the park after school	*The students sits in the park after school
The sisters love to run in the forest	*The sisters loves to run in the forest
The brothers attend football practice every day	The brothers attends football practice every day

3) Main clauses with long distance agreement, singular subjects

The house with yellow and white doors looks nice	*The house with yellow and white doors look nice
The teacher with black shoes walks to work every day	*The teacher with black shoes walk to work every day
The boy with blue eyes seems very happy	*The boy with blue eyes seem very happy
The girl with golden earrings takes the bus to school	*The girl with golden earrings take the bus to school
The boy with broken arms tries to read a book	*The boy with broken arms try to read a book
The book about fast cars makes the girl happy	*The book about fast cars make the girl happy

4) Main clauses with long distance agreement, plural subjects

The boys in the black car look very scary	*The boys in the black car looks very scary
The parents with the nice car talk to their kids	*The parents with the nice car talks to their kids
The girls with short blonde hair like to sit in the park	* The girls with short blonde hair likes to sit in the park
The cats with long white fur drink milk every day	*The cats with long white fur drinks milk every day
Those tourists with the heavy suitcase seem tired	*Those tourists with the heavy suitcase seems tired
The kids with the red bike play in the garden	*The kids with the red bike plays in the garden

Appendix 6: Filler sentence pairs

My father went to the shop every monday	*Father my went to the shop every monday
Very soon the band will their favourite song	*Very soon will the band play their favourite song
The children played the guitar every day	*Children the played the guitar very well
I often cook dinner after work	*I cook often dinner after work
Death can be scary	*Death be scary can
Sometimes Peter wrote letters	*Peter wrote sometimes letters
I usually play football on the weekends	*I play usually football on the weekends
The black shoes are too small for the woman	*The black shoes are small for woman the

We usually eat eggs for breakfast	*We eat usually eggs for breakfast
The girl ate a lot	*Girl the ate a lot
Hope will keep us alive	*Hope keep us alive will
Life can be difficult	*Life be difficult can
Thomas is a strict teacher	*Thomas is teacher a strict
Next week the students must practice for their exam	*Next week must the students practice for their exam
My mother read a book about flowers	*Mother my read a book about flowers
Nina often sang songs	*Nina sang often songs

*Susan always drank milk	*Susan drank always milk
Jack and Ruth walk to church every Sunday	*Jack and Ruth walks to church every Sunday
We drink coffee in the morning sometimes	*We drink sometimes coffee in the morning
John and Jane drive to work every day	*John and Jane drives to work every day
Susan hit him for being mean	*Susan for being mean hit him
Emma saw him at school	*Emma him saw at school
Mark runs in the park every day	*Mark every day runs in the park
Peter is a bad student	*Peter a bad student is

Appendix 7: List of subjects and verbs used in the practice part of the PPP session

The man	Walks
The girl	Runs
The boy	Swims
The woman	Juggles
The Clown	Drives
The dog	Jumps
The dragon	Supports
The cat	Spends

The player	Demands
The business man	Kicks
The bus driver	Turns
The chef	Holds
The tiger	Drinks
The wolf	Sleeps
The carpenter	Greets
The dentist	Takes

The men	Walk
The girls	Run
The boys	Swim
The women	Juggle
The clowns	Drive
The dogs	Jump
The dragons	Support
The cats	Spend

The players	Demand
The business men	Kick
The bus drivers	Turn
The chefs	Hold
The tigers	Drink
The wolves	Sleep
The carpenters	Greet
The dentists	Take

Appendix 8: Questionnaire

Language background:

Spørreskjema – bakgrunnsopplysninger:

Kjønn	**Fødselsår**
Mann☐ Kvinne☐	_____

Hvilke språk bruker dere hjemme?

_____

Har du gått på skole i Norge fra første trinn? Ja Nei

Hvis nei, i hvilket trinn begynte du i Norge? _____

Hvilke andre språk har du hatt på skolen eller kurs?

_____

Bruker du et annet språk enn norsk i ditt daglige liv (utenfor skolen)? Ja Nei

Hvis ja, hvilke språk? _____

Hva bruker du fremmedspråket til? _____

_____

Hvor gammel var du da du startet å snakke Engelsk? _____