UiT The Arctic University of Norway

Faculty of Humanities, Social Sciences and Education

Cross-linguistic Influence in L3 Acquisition of English by Persian Heritage Speakers in Norway

Farnoosh Abdollahi Dehooei

Master's thesis in English Linguistics, ENG-3991, September 2022

Dedication

To my mother for her unconditional love and support

Acknowledgement

First and foremost, I am deeply grateful to my supervisor Prof. Marit Westergaard for her patience, support, invaluable comments, and constructive feedback through all steps of my writing process, without whom this thesis would have not been possible. I would also like to thank my co-supervisor Dr. Khadij Gharibi for her valuable guidance and advice.

Similarly, I would like to express my gratitude to Dr. Sergey Minor, Prof. Bjørn Lundquist, and Prof. Natalia Mitrofanova who patiently helped me with the statistical analyses. I would also like to thank Dr. Nadine Kolb and Dr. Aleksandra Tomić for sharing their knowledge and experience with me in designing my study.

Many thanks to all the participants who took part in my pilot studies, and also all the Iranian and Norwegian families who contributed to this thesis. I extend my special thanks to my former colleagues in Iran, who helped me a lot with data collection, and also my friend, Melika Rajabi, who was always there when I needed help.

I express my heartfelt gratitude to my husband, Hooman, for his endless love and support, for believing in me the whole time, even in my most desperate moments. Last, but not least, I would like to express my deepest gratitude to my mother, Soghra, for all the sacrifices she made for my progress.

Abstract

Aims and objectives: This thesis investigates cross-linguistic influence (CLI) in child third language (L3) acquisition. The main objectives are to find the source of CLI, whether it is L1, L2, or both, and find out whether CLI occurs on a property-by-property basis or as a wholesale phenomenon.

Methodology: Persian-Norwegian heritage bilinguals acquiring L3 English are compared to L2 English learners with either L1 Persian or L1 Norwegian (matched in English proficiency). The participants were tested through a self-paced reading task and an acceptability judgment task in five language properties: adjective placement (Adj-N), definiteness and gender (where English is similar to Norwegian) subject-verb word order in non-subject-initial declaratives (SV), and adverb-verb (Adv-V) word order in subject-initial declaratives (where English is similar to Persian)

Data and analysis: A total of 82 participants were tested in English (L3 learners = 29, L2 learners with L1 Norwegian = 29, L2 learners with L1 Persian = 24). The L3 learners were also tested in their background languages (Persian and Norwegian).

Findings: The findings point to a major role of the dominant language (Norwegian), and the analyses showed that age and English proficiency are two significant factors. There was facilitative and non-facilitative CLI from Norwegian in the L3 group. Therefore, the findings do not go against wholesale CLI. While the self-paced reading task did not show any effects of grammaticality or group, there were two critical properties in the acceptability judgement task, definiteness and SV. In definiteness, the L3 group and the Norwegian controls were statistically similar, and they outperformed the Persian control group. In the SV property, however, the Persian controls outperformed the other two groups, while the Norwegian controls and the L3 group were statistically the same again.

Significance: This study adds to the existing data on child L3 acquisition and contributes to the ongoing discussions on the role of previously acquired languages, mainly the source and nature of CLI.

Keywords: Heritage speakers, Cross-linguistic influence, Third language acquisition, LPM, TPM.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

3	third person
ACC	accusative
Adj-N	adjective-noun word order
Adv-V	adverb-verb word order
AJT	Acceptability judgment task
ASP	aspectual
CEM	cumulative enhancement model
CLI	cross-linguistic influence
DEF	definite
Eng	English
EZ	ezafe
INDF	indefinite
L2	second language
L2A	second language acquisition
L2SF	second language status factor
L3	third language
L3A	third language acquisition
LPM	linguistic proximity model
Nor	Norwegian
Per	Persian
PL	plural
POSS	possessive
PRS	present
RT	reaction time
SG	singular
SPR	self-paced reading task
SV	subject-verb word order
TPM	typological primacy model
V2	verb second word order

1 Introduction

The field of third language acquisition (L3A) has gained great interest in recent years, and a growing number of research studies have explored the acquisition of an L3 across different language combinations using different methodologies (e.g., Bardel et al., 2007; Hermas, 2015; Jin, 2009; Flynn et al., 2004; Rothman, 2011, 2015; Westergaard et al., 2017; Lago et al., 2019).

One of the key questions in the field of L3A regards the source of cross-linguistic influence (CLI). Unlike second language acquisition (L2A), where there is only one source of CLI (L1), there are two potential sources of CLI in L3A. Therefore, the question is whether both of the previously acquired languages influence L3A, or whether there is only one source of CLI. If both previously learned languages affect L3A, it is assumed that CLI source selection is a property-by-property process (Westergaard et al., 2017; Slabakova, 2017; Westergaard, 2019, 2021a). However, if the CLI is solely from a language that is typologically similar to the target language, CLI is assumed to occur in one fell swoop (Rothman 2011, 2015). Various factors have been argued to affect CLI source selection, such as order of acquisition (Bardel et al., 2007; Jin, 2009; Hermas, 2010, 2015), language dominance (Fallah et al., 2016), overall typological similarity (Rothman et al., 2010) and structural proximity (Westergaard et al., 2017). Although many research studies have found empirical evidence for CLI source and nature, it is still the topic of many debates due to divergent results across studies.

In light of this, the main objectives of this study are twofold: To find out 1) whether CLI is only from one of the previously learned languages or both, and 2) whether CLI is a wholesale or property-by-property phenomenon. Persian-Norwegian heritage speakers acquiring L3 English are investigated in order to answer these research questions. The participants are tested in five language properties, including adjective placement, subject-verb word order in non-subject-initial declaratives, adverb placement is subject-initial declaratives, gender, and definiteness. English patterns with Norwegian in three of the properties (adjective placement, definiteness, and gender), and it is similar to Persian in the other two properties (subject-verb word order in non-subject-initial declaratives and adverb placement in subject-initial declaratives).

Persian-Norwegian heritage speakers were chosen as the target population, since the background languages (Persian and Norwegian) are very different from each other, one of them is superficially similar to the L3 (English), while the target language shares properties with both. The heritage speakers (L3 group) are compared to Norwegian and Persian L2 learners of English. The groups are matched in English proficiency. All participants are tested through an Acceptability Judgment Task (AJT) and a Self-Paced Reading (SPR) task. In addition to the AJT and SPR, the L3 group is tested in the background languages (Norwegian and Persian) through an audio AJT in order to make sure that they have acquired the properties under investigation in Norwegian and Persian.

The SPR task shows no statistically significant results, and a ceiling effect is observed in some of the properties (e.g., Adj-N property) in the AJT, but significant differences are found in two properties (definiteness and SV) in the AJT. The results point to facilitative and non-facilitative CLI from Norwegian in the definiteness and SV properties respectively. The findings of the present study are argued to be best explained by the role of the dominant language, since the L3 learners have lived in Norway their whole life, where the majority language is Norwegian, and they have learned English in Norwegian classrooms, through the medium of Norwegian. For the question of wholesale vs. property-by-property CLI, the findings are inconclusive.

Answering the questions addressed in the present study contributes to the ongoing debates on the role of previously acquired languages in L3A and also adds to the existing data on child L3A.

This thesis is structured as follows: the next chapter introduces the theoretical background, including the models of L3A and empirical evidence, and discusses cross-linguistic variation in English, Norwegian, and Persian. In chapter 3, the design of the study and research questions are outlined as well as predictions. Additionally, the results of several pilot studies, which were conducted before the main study, are described. In chapter 4, the results of this study are presented along with the statistical analysis, and the findings are discussed in chapter 5. Chapter 6 is a brief summary that concludes this thesis.

2 Background

In this chapter, I first discuss a core term in this study, cross-linguistic influence, and then describe the characteristics of the target population, heritage speakers. Next, I provide an overview of the literature in third language acquisition and the existing models and their assumptions, mainly the two dominant models, the Typological Primacy Model (TPM) and the Linguistic Proximity Model (LPM). Finally, I detail the similarities and differences between the languages involved in the present study with respect to the properties under investigation – gender, definiteness, adjective-noun word order, verb second word order in non-subject-initial declaratives, and adverb-verb word order in subject-initial declaratives.

2.1 Cross-linguistic Influence

Grammar is a collection of linguistic representations. Sharwood Smith (2019, p.10) defines representation as a "network of features expressing some basic structural category", to be more precise, a linguistic representation includes the syntax, morphology, phonology, semantics, etc., of a given language in the mind of a speaker. The grammar then consists of the aforementioned language domains which is ultimately utilized to parse linguistic input for the purpose of comprehending or producing a particular language. Consider a bi- or multilingual mind in which there are two or more linguistic representations. Cross-linguistic influence (CLI) is defined as the influence that one language system has on another language system in the mind/brain, which may occur at the lexical, structural, and phonological levels and affect the usage and processing of a language (Sharwood Smith, 1983, 1989; Kellerman & Sharwood Smith, 1986).

The terms CLI and transfer are often used interchangeably in the L3 literature. However, Sharwood Smith (2021, p.410) argues that, although transfer is a "handy metaphor", it is a misleading term that has limitations in describing this phenomenon. He continues that transfer indicates that there is movement of grammatical features from one language to another, and as a result, the host grammar will deteriorate, which does not actually happen. Since the term transfer implies switching from one place into another, it is a misleading terminology to describe the process.

Therefore, throughout this study, the term CLI will be used since it seems to be a more sufficient term than *transfer* and covers different forms of interference or interactions at both

"product and process" levels (Sharwood Smith & Truscott, 2014, p.194). The term *transfer* is only used to describe and discuss the models and the previous works that use this term in this study.

CLI may be facilitative or non-facilitative. Facilitative (positive) CLI occurs when a property in one or both background languages has significant similarity with the corresponding structure in the target language. As a result, the speaker is able to parse the input successfully and produce grammatical speech in the target language. On the other hand, non-facilitative (negative) CLI occurs when a structure in one or both previously acquired languages is different from the corresponding structure in the target language. The speaker may then parse the input incorrectly which typically leads to difficulties in comprehension. Another scenario for non-facilitative CLI would be when the speaker has not had sufficient L3 input which would typically lead to producing ungrammatical speech in the target language (see Westergaard, 2021b for more details).

2.2 Heritage Speakers

Since the focus of the present study is on heritage speakers, it is necessary to explain what is meant by the term and who they are. There are numerous characterizations mentioned for heritage speakers in the literature (Montrul, 2008; Rothman, 2009; Polinsky, 2018; Lohndal et al., 2019). All of them, despite their differences, have some points in common:

i. Heritage speakers are bilingual.
ii. Heritage speakers speak a minority language (heritage language) in an environment with a majority language.
iii. When the heritage speakers reach adulthood, they are typically more proficient in the majority language of the environment.

Heritage speakers are not a homogeneous population, and they differ in their level of proficiency; some of them are relatively balanced bilinguals, while some may have a lower proficiency in their heritage language, mostly due to social factors and language experiences. Heritage speakers normally have considerable exposure to the heritage language during childhood at home from their caretakers. The exposure typically decreases as the heritage speaker grows up and enters the society of the majority language. However, it is possible for the exposure to increase in adulthood in cases where the speaker moves to a place where the

heritage language is spoken as a majority language. Therefore, heritage speakers are a heterogenous group of speakers with varying proficiency in the heritage language. The heritage language spoken by the heritage speakers is not always the same as the language spoken in the homeland.

In the present study, the term 'heritage speaker' is used to refer to an individual who has some level of proficiency in the minority/heritage language and speaks the language regularly at home or in other national environments and has at least one parent who spoke/speaks the heritage language to him/her as the first language in childhood. Most of the heritage speakers in the present study were simultaneous bilinguals who were exposed to both Persian and Norwegian from birth or started learning Norwegian a bit later. Furthermore, they were able to speak and understand Persian, but they could not read or write in Persian.

2.3 L3 Models and Previous Studies

Third language acquisition (L3A) is a relatively new field, and thus most of the work done on this topic build on previous studies conducted within L2 acquisition, especially works regarding CLI. However, L3A is a more complex process than L2 acquisition: in L2 acquisition there is only one source of influence, the L1, while in L3A there are two language systems available to the learner. In the last decade (or so), several models of L3A have been proposed. The main research questions in these models concerned the source of CLI, i.e., whether the L1, the L2 or both are the source of influence, whether CLI is wholesale or property by property, and also the effects of other variables such as the role of the language of communication, typological or structural similarity etc.

2.3.1 The Default L1 Effect

One possible scenario for the source of influence in L3A is L1 influence. Although no model of L1 influence in L3A has been proposed, there are some studies that have found strong L1 influence (Jin, 2009; Na Ranong & Leung, 2009; Hermas, 2010, 2015), suggesting that the native language might be the default source of influence in L3A, since the learners may be more proficient in their L1, making it more accessible for transfer (see Lloyd-Smith et al., 2017).

One of the studies that provided evidence for L1 influence is Jin (2009). This study set out to investigate the role of the L1 and L2 in L3 acquisition of Norwegian objects by Chinese graduate students in Norway who were advanced learners of L2 English. Norwegian and

English are both subject-prominent languages, while Chinese is a topic-prominent language; that is to say, Chinese allows null objects, while a referential pronoun or noun phrase is necessary in Norwegian and English. The results of grammaticality judgment and sentence correction task showed variability in the rejection of null objects in the L2 and L3. Group results showed that Chinese learners rejected English null objects with high accuracy (70%). It was further shown that over half of the individuals judged and corrected English null objects to a native-like level. However, Chinese learners had difficulty in rejecting the null object sentences in Norwegian, which was an indication of non-facilitation from the L1 (Chinese). Overall, the researchers concluded that their results show a strong negative effect of L1 Chinese and little influence of L2 English in L3 acquisition of Norwegian objects.

Furthermore, Hermas (2010) studied the L3 acquisition of verb movement by adult Arabic-French bilinguals who were beginner learners of L3 English. The participants' knowledge of this parameter was tested through an acceptability judgment task and a preference test. French and Arabic are similar since they display verb movement, whereas English lacks this property. Adverbs only follow the verbs in French. In Arabic, adverbs may precede or follow the verbs, and adverbs are placed preverbally in English. The results of both tasks revealed that the L3 learners' accuracy in English was negatively influenced only by their L1 Arabic, and the L3 group differed significantly from the French and English natives. The results were interpreted as evidence for the special status of the L1 as the source of influence in the initial states of L3 English.

2.3.2 The L2 Status Factor

Contrary to the default L1 scenario, the L2 Status Factor (L2SF) hypothesis proposes that the L2 has a privileged role to be the source of influence in L3 acquisition (Bardel & Falk, 2007, 2012; Falk & Bardel, 2011). The L2SF hypothesis is based on the Declarative/ Procedural model by Paradis (2009), which proposes that native and non-native grammars are stored in different places in the mind/brain, the former in procedural memory and the latter in declarative memory. Therefore, since an L3 is acquired in the same way as an L2, both the L2 and L3 are stored in declarative memory, which makes transfer from L2 to L3 easier than L1 to L3.

Bardel and Falk (2007) investigated the acquisition of negation in either L3 Dutch or L3 Swedish by adult speakers who had different L1s and L2s. The focus of the study was the

V2 phenomenon. Among the previously learned languages in this study, German, Dutch, and Swedish are V2 languages, while Albanian, Italian and English are not. The researchers made sure that the participants had one V2 language and one non-V2 language as either L1 or L2. Ten sessions of Swedish lessons were video and audio recorded. After examining the oral data, the findings showed that the participants with a V2 language as their L2 performed significantly better in acquiring either Dutch or Swedish than the ones who had a V2 language as their L1. The researchers concluded that the L2 acts like a filter that blocks L1 influence in L3 acquisition.

2.3.3 The Cumulative Enhancement Model

The Cumulative Enhancement Model (CEM) (Flynn et al., 2004) marks a shift in understanding the role of previously learned languages in L3A. In contrast to the Default L1 Effect and L2SF, the CEM argues that influence in L3A is selective and can come from any source as long as it is facilitative. This means that the source of influence can be L1, L2 or both, and that previously learned languages should affect the L3A either positively or have no effect. Thus, a learner first examines a particular property in the background languages; if there is an equivalent for that property in one of the languages, it is then selected to transfer; in case of no equivalent, the property is learned. The CEM is based on work by Flynn et al. (2004), in which they investigated the role of L1 and L2 in acquisition of English restrictive relative clauses through an elicited imitation task in three groups of participants: L1 Kazakh-L2 Russian learners of L3 English, L1 Spanish and L1 Japanese learners of L2 English. English, Spanish, and Russian are similar in that they are all head-initial languages, while Kazakh and Japanese are head-final. The results showed that the bilingual group (L1 Kazakh-L2 Russian) and the L1 Spanish group performed similarly, while the L1 Japanese group behaved differently. From this, the authors concluded that since the bilingual and the L1 Spanish groups had already acquired a language (L1 or L2) with the head-initial parameter, they were able to perform better than the L1 Japanese group that had the head-final parameter. In general, the findings suggested that all previously acquired languages can positively influence the acquisition of a third language, which led the authors to propose the CEM.

2.3.4 Typological Primacy Model

The next model in the field of L3A is the Typological Primacy Model (TPM; Rothman 2011, 2015; Rothman et al., 2010; Rothman et al., 2019).

According to the TPM, the learner copies the whole linguistic system of the language that is typologically closer to the target language, which then constitutes the initial grammar of the L3[1]. The TPM argues that the overall typological similarity between the target language and one of the previously learned languages determines the source of the influence. Therefore, the TPM is similar to the CEM in that the source of the influence can be L1 or L2, but crucially not both. However, unlike the default L1 effect and L2SF in which order of acquisition matters, overall typological similarity determines the source of influence in this model, no matter whether the influence is facilitative or non-facilitative. There are supporting evidence from empirical studies, where results were interpreted as transfer from the typologically related language (e.g., Hopp, 2019). Hopp (2019) investigated child L3A in Turkish-German heritage speakers learning English compared to German monolinguals learning English. The results from sentence repetition and oral sentence production tasks showed transfer from German in both groups, and therefore they were interpreted as evidence for transfer from the typologically closer language to the target language (English). However, Hopp (2019) pointed out that the L3 learners were more dominant in German than Turkish, and this may affect the transfer patterns. Additionally, transfer from German could be because English was taught via the majority language German, and students may increasingly note similarities between German and English (Hopp et al., 2019).

Rothman (2015, p.184) explains further that wholesale transfer is practical from a cognitive economy perspective, because when transfer occurs in one fell swoop, each property does not have to be contrasted against two highly activated languages. Furthermore, among other executive control system that are required in a bilingual mind, inhibition is needed to restrain the activation of other languages. Thus, wholesale transfer based on overall typological similarity between the target language and other grammars would be more efficient. However, more recent works on the TPM (e.g., Rothman, 2015; Rothman et al., 2019) suggest that property-by-property transfer is in fact possible (before and after wholesale transfer), especially

[1] Wholesale transfer in the TPM is an extension of the Full Transfer/Full Access Model by Schwartz and Sprouse (1996, p.40) in L2 acquisition in which they argue "the initial state of L2 acquisition is the final state of L1 acquisition (Full Transfer) and the failure to assign a representation to input data will force subsequent restructurings, drawing from options of UG (Full Access)."

in L4 acquisition, because "L3 experiences of non-facilitation might very well mean that full transfer will be disregarded as a viable option when the mind is an experienced multilingual one." (Rothman et al., 2019, p.157). Therefore, structural similarity is a deterministic factor in assessing the typological similarity.

How does the parser determine typological similarity? Rothman (2013, p.238) has introduced a hierarchy of properties by which the parser determines the typological/structural similarity. This hierarchy is outlined below in order of relative impact/influence:

1. Lexicon
2. Phonology/Phonotactics
3. Functional Morphology
4. Syntactic Structure

This hierarchy of properties enables the parser to decide which property is the most similar to one of the previously acquired languages in order to select a source of influence. If similarity at one level is not sufficient for CLI to occur, then the parser will turn to the next level and so on.

2.3.5 Linguistic Proximity Model

The Linguistic Proximity Model (LPM; Westergaard et al., 2017; Westergaard, 2019) shares some similarities with the TPM and differs from the TPM in some respects. The LPM considers transfer as a property-by-property phenomenon that allows for facilitative as well as non-facilitative influence, and the source of influence can be either one of the previously acquired languages or both. Furthermore, this model assumes that CLI occurs when there is an abstract structural similarity between a linguistic property in the target language and properties in the background languages. Thus, the LPM is a structural model (similar to the TPM); however, it rejects the idea of wholesale transfer (a view also shared by the Scalpel Model[2] and other models discussed earlier except the TPM). Westergaard (2021b) explains further that, overall typological/ lexical similarity may override structural similarity at early stages.

[2] The Scalpel Model proposed by Slabakova (2017) shares many similarities with the LPM; they both assume transfer takes place property-by-property and that L3 acquisition is a cumulative process. In both models, structural similarity is a motivating factor for CLI. Due to these similarities, the Scalpel Model is not discussed in this study.

According to the LPM, all previous grammars are active and available to the learner in L3A; however, it is possible that " … the typological/lexical similarity between the L3 and one of the previously acquired languages will cause stronger activation of the syntactic structure of this particular language" (Westergaard, 2021b, p.6). Thus, facilitative influence is a result of structural similarity, and non-facilitation is the result of misanalysis of L3 input, which will lead the learner to produce structures that are not target-like. Furthermore, Westergaard et al. (2017, p.670) argue that property-by-property transfer is a more efficient process cognitively since "this would reduce the amount of effort required to unlearn incorrectly transferred properties". Westergaard (2019, p.393) elaborates further that one cannot say for sure if transferring the whole grammar once is cognitively more economical or transferring little many times.

Recent work on the LPM (Westergaard, 2019; 2021a; 2021b) has defined CLI as a result of co-activation of the previously acquired grammars (to different degrees based on structural similarity) in order to parse L3 input. During this process, the entirety of the previous grammars is available to the parser.

Thus, emphasizing property-by-property transfer, Full Transfer Potential (FTP) argues that "anything may transfer", not that "everything does transfer" (Westergaard, 2019, p.389). Therefore, the parser uses the properties of the previously acquired languages, which are activated simultaneously, to parse the L3 input to build a representation that is unstable and weak in the beginning, but further input and parsing will strengthen it.

In my view, the TPM and LPM are currently the two most promising and dominating models in the field of L3A. The Default L1 Effect and L2SF assume a privileged role for either L1 or L2 to be the sole source of CLI, and while there are studies which have provided evidence in support of these models, they cannot both be right.

The CEM does not account for non-facilitative influence, but there are numerous studies that provide evidence for non-facilitation. It was shown in a systematic review of 71 L3 studies by Puig-Mayenco et al. (2018) that the Default L1 Effect and L2SF account for only 14.1% and 28.2% of the data respectively. And the CEM accounts for only 5.9% of the data. Finally, the TPM is a model that is limited to the initial stages and does not make predictions for later stages,

while the LPM can make prediction for any stages of L3A. Thus, the LPM and to some degree the TPM will be included in the analysis of the present study.

2.3.6 The Dominant Language of Communication

There are other external factors that might affect the selection of source(s) of influence in L3A, e.g., the dominant language of communication. The language of communication is defined as the language that is more frequently spoken by the subjects in different contexts (Fallah et al., 2016, p.226). Fallah et al. (2016) investigated the initial stages[3] of L3A of English possessives by Mazandarani-Persian bilinguals. The analysis of the results showed that the dominant language of communication is a determining factor in CLI source(s) selection in initial stages of L3A.

However, the effect of language of communication will not be addressed in the present study, because while collecting data about the participants' language background, in-depth data about their language use were not collected.

2.4 Crosslinguistic Differences in Persian, Norwegian and English

English and Norwegian have close similarities since they are both Germanic languages, while Persian is a typologically distant language and belongs to the Indo-Iranian languages with a totally different script and writing system. However, these three languages behave differently in the language properties under investigation that is to say, there are some structural similarities between English and Norwegian, namely gender, definiteness, and adjective placement (where Persian is different), while English and Persian pattern together in subject-verb word order in non-subject-initial declaratives and adverb-verb word order in subject-initial declaratives. The reason behind selecting these particular properties is that they are considered to be problematic for Norwegian and Persian L2 learners of English respectively. There are previous research studies for some of the properties, showing that they cause difficulties for L2 learners of English. For example, definiteness is a problematic property for Persian L2 learners of English, even at advanced levels, since Persian lacks definiteness (Geranpayeh, 2000; Momenzade &

[3] As mentioned in Fallah et al. (2016, p.235), on average, the participants have had 24 hours of exposure to English. Then it could be argued that these participants are at the initial stages of L3A.

Youhanaee, 2014; Joolaee & Ghonsooly, 2015; Kargar, 2019). To my knowledge, the acquisition of adjective placement and gender by Persian L2 learners of English has not been studied before[4], and the reason why these two properties were selected is because they are different from English, and thus may be problematic for the Persian L2 learners of English. Furthermore, non-V2 word order seems to cause difficulties for Norwegian L2 learners of English in acquiring basic SVO word order, as their L2 English is often influenced by their L1 Norwegian in non-subject-initial declaratives and sentences with adverbs (Westergaard, 2003).

In the remainder of this section, the grammar of the above language properties in the three languages will be described in detail.

2.4.1 Gender

In the gender property, the focus is on personal pronouns. Gender plays a minor role in English grammar and is expressed by only pronouns. In the case of personal pronouns, a distinction must be made in third person singular between masculine '*he*', feminine '*she*' and neuter '*it*'. Gender assignment in English has a very simple rule: nouns denoting male humans are masculine, nouns denoting female humans are feminine, and other nouns are neuter (Comrie, 1999, p.458). Table 1 shows some examples of English personal pronouns.

Table 1 - Third Person Personal Pronouns in English.

Masculine	My brother is a student. **He** studies math.
Feminine	My sister is a student. **She** studies math.
Neuter	We have a big table. **It** is very old.

The gender system in European languages like Norwegian is typically more complex. Norwegian has a three-gender system that makes distinctions between masculine, feminine, and neuter. Note that this is nominal gender, which is not semantically based. Gender is expressed not only on personal and possessive pronouns, but also on adjectives, articles, and demonstratives (Rodina & Westergaard, 2015, 2017). However, the focus of this study is only on personal pronouns. Similar to English, Norwegian distinguishes between the masculine

[4] There was a simultaneous master's project, in which the acquisition of Adj-N and gender in Persian L2 learners of English was investigated (Rajabi, 2022). The results will be discussed in the discussion chapter.

‘*han*’ (he), feminine ‘*hun*’ (she) and neuter ‘*det*’ (it) as examples in Table 2. In addition to ‘*det*’, there is the personal pronoun ‘*den*’, which has the same meaning but refers to masculine and feminine nouns. It should be noted that in Norwegian Bokmål both ‘*han’* and ‘*hun’* only refer to animate nouns, while ‘*det’* and ‘*den’* refer to non-animate nouns.

Table 2 - Third Person Personal Pronouns in Norwegian.

Masculine	Broren min er student. **Han** studerer matematikk. brother.DEF my is student **he** studies math ‘My brother is a student. He studies math.’
Feminine	Søsteren min er student. **Hun** studerer matematikk. sister.DEF my is student **she** studies math ‘My sister is a student. She studies math.’
Neuter	Vi har et stort bord. **Det** er veldig gammelt. we have a big table **it** is very old ‘We have a big table. It is very old.’

Unlike English and Norwegian, Persian has no gender. The pronoun ‘*ū*’ serves as an equivalent for both ‘*he*’ and ‘*she*’, and ‘*ān*’ is translated as ‘*it*’ in English, as in examples [1a] and [1b]

[1] a. khāhære mæn dāneshjū ast. **Ū** rīāzi mīkhānæd. (Persian)
Sister.EZ my student is.3SG **s/he** math ASP.study.3SG
‘My sister is a student. She studies math.’

b. mā yek mīze bozorg dārīm. **ān** kheylī ghædīmī æst.
we one table.EZ big have.1PL **it** very old is.3SG
‘We have a big table. It is very old.’

2.4.2 Definiteness

English and Norwegian mark definiteness on nouns since they both have a fully grammaticalized article system. English displays definiteness on singular and plural count and mass nouns by adding the definite article ‘*the’* before nouns (Aarts, 2011). Definiteness on a singular count noun is shown in example [2a], while example [2b] displays definiteness on a

plural count noun, and [2c] shows definiteness on a mass noun. As can be seen, marking definiteness is the same across these three examples with different nouns.

[2] a. Sarah has a bag. **The** bag is blue. (English)

b. Sarah has a lot of toys. **The** toys are over there.

c. Police thought **the** evidence was not enough.

Norwegian marks definiteness with three different suffixes that correspond to the gender of the nouns: '*-a'* for feminine singular nouns, '*-en'* for masculine singular nouns and '*-et'* for neuter singular nouns. Additionally, definiteness on plural nouns in Norwegian is marked by adding the suffixes '*-ene*' or only '*-ne*'. There is no gender opposition in plural nouns in Norwegian (Enger & Corbett, 2012). Table 3 provides some examples of definiteness in Norwegian.

Table 3 - Definiteness in Norwegian.

Singular	Masculine	Sara har en sekk. Sekk**en** er blå. Sarah has/have INDF bag bag.**DEF** is blue 'Sarah has a bag. The bag is blue.'
	Feminine	Har du en adresse? Ja, adress**en** er Dramsvegen. have you INDF address yes address.**DEF** is/are Dramsvegen 'Do you have an address? Yes, the address is Dramsvegen.'
	Neuter	Et tog stopper på sentralstasjonen. Tog**et** drar snart igjen. INDF train stops at central station train.**DEF** travels soon again 'A train stops at the central station. The train will leave again soon.'
Plural		Sara har mange leker. Leke**ne** er der borte. Sarah has/have many toys.INDF toys.**DEF** is/are there away 'Sarah has a lot of toys. The toys are over there.'

In contrast to English and Norwegian, Persian lacks a fully grammaticalized article system . As a result of that there is no overt article or morphological inflection in Persian corresponding to English '*the*' denoting a definite context (Faghih, 1997; Karimi, 1999; Geranpayeh, 2000; Ghomeshi, 2003; Rezai & Jabbari, 2010; Momenzade & Youhanaee, 2014; Kargar, 2019). However, there are two indefinite markers in Persian: suffix '*-ī*' and numeral

'*yek*', both meaning one, which are exemplified in [3a] and [3b]. According to Faghih (1997, p.137) and Afzali (2012, p.20), some Persian grammarians believe that the absence of indefinite markers is actually a definite marker, that is to say, bare nouns in Persian are definite as in [4a] and [4b]. Although Persian does not have a special overt definite marker, there are some means of expressing semantic features of specificity through the object marker '*-ra*', the Ezafe particle (-e)[5], demonstratives and possessives. In other words, Persian marks nouns for specificity, but not for definiteness. Although specificity is not the concern of this study, two examples are provided in [5a] and [5b] .

[3] a. sārā kīf**ī** dāræd. (Persian)
Sarah bag.**INDF** have.3SG
'Sarah has a bag.'

b. tīnā **yek** ketāb mīkhānæd.
Tina **INDF** book read.3SG
'Tina reads a book.'

[4] a. moælemhā khændīdænd.
teachers laughed.3PL
'The teachers laughed.'

b. gorbe oftād.
cat fell.3SG
'The cat fell.'

[5] a. mæn īn ketābhā rā khāndæm.
I this books ACC read.1SG
'I read these books.'

b. līvānæm shekæst.

[5] Ezafe is an unstressed vowel that is spelled as '*-e*' at the end of some words and links elements of a single constituent together (Ghomeishi, 1997, p.729).

glass.POSS broke.3SG

'My glass broke.'

2.4.3 Adjective Placement

English and Norwegian are similar regarding adjective placement (Adj-N): adjectives are placed pre-nominally resulting in Adjective – Noun word order in both languages. In English, adjectives have the same form for definite, indefinite, singular and plural nouns, as in examples [6a-d].

[6] a. A **black** cat (English)
Adj N

b. **Black** cats
Adj N

c. The **black** cat
Adj N

d. The **black** cats
Adj N

In Norwegian, adjectives don't maintain the same form for all nouns. Adjectives agree with the gender, number, and definiteness of nouns they modify, as shown in Table 4.

Table 4 - Adjective Placement in Norwegian.

		Masculine	Feminine	Neuter
Indefinite	Singular	En fin katt INDF nice cat 'a nice cat'	ei fin bok INDF nice book 'a nice book'	et fint hus INDF nice house 'a nice house'
	Plural	fine Katter nice.PL cat.PL 'nice cats'	fine bøker nice.PL book.PL 'nice books'	fine hus nice.PL house.PL 'nice houses'
Definite	Singular	den fine katten DEF.SG nice.DEF cat.DEF 'the nice cat'	den fine boka DEF.SG nice.DEF book.DEF 'the nice book'	det fine huset DEF.SG nice.DEF house.DEF 'the nice house'
	Plural	de fine kattene DEF.PL nice.DEF cat.DEF 'the nice cats'	de fine bøkene DEF.PL nice.DEF book.DEF 'the nice books'	de fine husene DEF.PL nice.DEF house.DEF 'the nice houses'

Persian behaves differently from English and Norwegian in that adjectives are post-nominal, resulting in Noun – Adjective word order. Adjectives maintain their original form

across different types of nouns (singular or plural, definite or indefinite), and nouns are linked to adjectives by adding Ezafe (*-e*) as can be seen in example [7a-c].

[7] a. yek ketābe bozorg
INDF book.EZ big
'a big book'

b. ketābe bozorg
book.EZ big
'the big book'

c. ketābhāye bozorg
books.EZ big
'the big books'

2.4.4 V2 Word Order

A common feature of most Germanic languages is verb second (V2) word order, which is standardly considered to be the result of moving the finite verb to the second position in main clauses. An exception is modern English which has SVO word order. V2 word order is investigated in two types of declarative clauses in this study: non-subject-initial declaratives and subject-initial declaratives with adverbials.

2.4.4.1 SV Word Order in Non-subject-initial Declaratives

In Norwegian, a V2 language, it is necessary to have the finite verb in the second position. Therefore, the word order in non-subject-initial declaratives in Norwegian is XVS, as in [8]. In English, however, XVS word order in non-subject-initial declarative sentences is generally ungrammatical[6]. Therefore, even when the sentence does not begin with a subject, English maintains SV word order in non-subject-initial declaratives, as in [9]. As for Persian, although the word order in spoken Persian is quite flexible, the written language has a rigid

[6] Although English is normally not considered a V2 language, it is assumed that English has lost its V2 property during the Middle English period, but there still exist some cases of V2 in English, limited to certain clause types and verb types. Therefore, it is argued that English should be considered as a mixed V2 language (for a detailed overview, see e.g., Westergaard, 2007).

SOV word order. Therefore, transitive verbs are avoided in non-subject initial declaratives in this study, so that both Persian and English have XSV word order as shown in the examples below.

[8] I forrige uke **døde bestemoren min**. (Norwegian)
X V S
in last week **died grandma.DEF POSS**
'Last week my grandma died.'

[9] Last week **my grandma died**. (English)
X S V

[10] hæfte pīsh **mādærbozorgæm mord**. (Persian)
X S V
Week last **grandma.POSS died.3SG**
'Last week my grandma died.'

2.4.4.2 Adv-V in Subject-initial Declaratives with Adverbs

Since Norwegian is a V2 language, it exhibits V-Adv word order in subject-initial declaratives with frequency adverbs. In English, adverbs are normally placed before lexical verbs, resulting in Adv-V word order. Similar to the previous condition, Persian patterns with English, where frequency adverbs are placed before verbs (Adv-V).

[11] Broren min **sover alltid**. (Norwegian)
V Adv
Brother-DEF my **sleep.PRS always**
'My brother always sleeps.'

[12] My brother **always sleeps**. (English)
Adv V

[13] bærādæræm **hæmīshe mīkhābæd**. (Persian)
Adv V
Brother.POSS **always ASP.sleep.3SG**
'My brother always sleeps.'

2.5 Chapter Summary

This chapter first introduced the key concept of this study – cross-linguistic influence, a phenomenon in which the grammar of one language influences another language – and then described the existing models and theories in this field. These models focused specifically on variables that can determine the source of CLI.

The Default L1 Effect and the L2SF argue that order of acquisition is a determining factor in selecting CLI source. The Default L1 Effect assumes that the first language has a privileged role in affecting the target language, while L2SF argues that the L2 is a more accessible source for CLI. There is empirical evidence for both of these models that support the fact that the source of CLI can be either L1 or L2. However, the two dominant models discussed here assume that order of acquisition is not as significant as other variables. The TPM introduces a hierarchy of properties, based on which the parser selects a language (L1 or L2) which is typologically more similar to the target language to be the sole source of CLI. The LPM assumes that CLI occurs as the result of co-activation of the previous grammars, and CLI source selection is based on structural similarity between a particular language property in the previous grammars and a property in the target language.

The syntactic differences between the three languages in the triad, Persian, Norwegian and English, were then discussed in detail. While Norwegian and English are very similar languages and pattern together in gender, definiteness, and adjective placement, they differ in SV and Adv-V word order (where Persian patters with English).

Table 5 - Summary of the Properties

Language Properties	
Gender	**(Eng = Nor ≠ Per)**
Definiteness	**(Eng = Nor ≠ Per)**
SV in non-subject initial declaratives	**(Eng = Per ≠ Nor)**
Adv-V in subject initial declaratives	**(Eng = Per ≠ Nor)**
Adj-N	**(Eng = Nor ≠ Per)**

In the following chapter, I introduce the research questions and hypotheses which are based on the models discussed here. I then explain the design of the study and provide details about the methods used and also briefly explain the pilot study.

3 Research Questions and Methodology

In this chapter the objectives of the present study are introduced, and then predictions are made based on the models and supporting evidence, which were introduced in the previous chapter. The rest of this chapter is allocated to the methodology– self-paced reading and acceptability judgement task. I then describe the test items, procedure, and the pilot studies are described and discussed at the end. This study has been approved by the Norwegian Centre of Research Data (NSD).

3.1 Research Questions and Predictions

The research questions addressed in this study are as follows:

RQ 1. What is the source of CLI? Is it only one of the previously learned languages or both?

RQ 2. Does CLI occur property by property or as a wholesale phenomenon?

Following the assumptions of the LPM, it is hypothesized that: 1) the source of (facilitative and non-facilitative) CLI will be from both previously acquired languages, and 2) CLI occurs on a property-by-property basis. Moreover, the TPM is a model exclusive to the initial stages and does not make predictions for later stages of acquisition except as a direct development of wholesale CLI. The reason the TPM is considered in the present study is that, in the field of L3A, only the TPM argues for wholesale CLI. Therefore, it is hypothesized based on the TPM that: 1) Norwegian will be the sole source of (facilitative and non-facilitative) CLI, and 2) CLI will occur as a wholesale phenomenon. The following predictions are made in the present study if CLI occurs property-by-property according to the LPM:

1. In the self-paced reading task: For language properties where Norwegian and English pattern together, but Persian is different, it is expected that there will be a significant difference between the RTs of the grammatical and ungrammatical sentences in the L1-Norwegian group. The L1-Persian group will not have a significant difference between the grammatical and ungrammatical conditions in these properties. The RTs of the L3 group are expected to be in the middle of the two control groups in both grammatical and ungrammatical condition.

2. In the self-paced reading task: For language properties where Persian and English pattern together, but Norwegian is different, it is expected that there will be a significant difference between the RTs of the grammatical and ungrammatical sentences in the L1-Persian group. The L1-Norwegian group will not have a significant difference between the grammatical and ungrammatical conditions in these properties. The RTs of the L3 group are expected to be in the middle of the two control groups in both grammatical and ungrammatical condition.
3. In the acceptability judgement task: For language properties where Norwegian and English pattern together, but Persian is different, it is expected that the L1-Norwegian group will score the highest and the L1-Persian group the lowest, while the L3 group is expected to be in the middle.
4. In the acceptability judgement task: For the language properties where Persian and English pattern together, but Norwegian is different, it is expected that the L1-Persian group will score the highest and the L1- Norwegian the lowest, while the L3 group is expected to be in the middle.

The reason the L3 group is expected to be in the middle is that, based on the LPM, both previously acquired languages are available as potential CLI source(s) to the parser. Therefore, the learner will receive both facilitative and non-facilitative influence from the background languages, which will in turn result in being in the middle of the two control groups. However, this outcome is expected only if the timing is right, so that floor and ceiling effects are avoided (see Westergaard et al., 2022 for more details). Facilitative influence occurs when the L3 and background languages pattern together regarding a particular property, while non-facilitative influence occurs when the L3 and the background languages are dissimilar regarding a particular property.

However, if wholesale CLI takes place, then the following predictions are made according to the TPM:

1. In the self-paced reading task: For language properties where Norwegian and English are similar, but Persian is different, there will be a significant difference between the RTs of the grammatical and ungrammatical sentences in the L3 and L1-Norwegian groups, while the L1-Persian group will not have a significant difference between the grammatical and ungrammatical conditions.

2. In the self-paced reading task: For language properties where Persian and English are similar, but Norwegian is different, there will be a significant difference between the RTs of the grammatical and ungrammatical sentences in the L3 and L1-Persian groups, while the L1-Norwegian group will not have a significant difference between the grammatical and ungrammatical conditions.
3. In the acceptability judgement task: For language properties where Norwegian and English are similar, but Persian is different, it is expected that the L3 and L1-Norwegian groups will score the highest (the difference between the two groups will not be significant) and the L1-Persian group the lowest.
4. In the acceptability judgement task: For the language properties where Persian and English pattern together, but Norwegian is different, it is expected that the L1-Persian group will score the highest and the L3 and L1-Norwegian the lowest (the difference between the two groups will not be significant).

The reason the L3 group is expected to perform like the L1-Norwegian group is that, according to the TPM and the hierarchy of properties introduced in the previous chapter, the parser will choose one of the previously acquired languages based on the highest level where the similarities are found. Thus, the learner is expected to select Norwegian as the sole source of CLI and copy the whole representation, whether the influence is facilitative or non-facilitative. Table 6 shows an overview of the predictions in the present study.

Table 6 - Overview of the Predictions.

Condition	LPM Predictions	TPM Predictions	
Gender	L1-Per[1] < L3 < L1-Nor[2]	L1-Per < L3 = L1-Nor	Eng[3] = Nor ≠ Per
Definiteness	L1-Per < L3 < L1-Nor	L1-Per < L3 = L1-Nor	Eng = Nor ≠ Per
Adj-N	L1-Per < L3 < L1-Nor	L1-Per < L3 = L1-Nor	Eng = Nor ≠ Per
SV	L1-Nor < L3 < L1-Per	L3 = L1-Nor < L1-Per	Eng = Per ≠ Nor
Adv-V	L1-Nor < L3 < L1-Per	L3 = L1-Nor < L1-Per	Eng = Per ≠ Nor

[1]Per = Persian, [2]Nor = Norwegian, [3]Eng = English

3.2 Method

3.2.1 Self-paced Reading Task

SPR is an online computerized method in which subjects read sentences that are divided into words or phrases (known as regions of interest) at their own pace by pressing a button. It is called *self-paced*, or sometimes *subject-paced*, because the participants decide how much time they need to spend reading each region, unlike fixed-paced methods such as Rapid Serial Visual Presentation (RSPV), in which the reading times are set before-hand by the researcher. Each time that the participants press a button (commonly the space bar), a word or phrase of the sentence appears, and a reaction time (RT, the time elapsed between each button press) in milliseconds (ms) is recorded. This method was first used by Aaronson and Scarborough (1976) to investigate first language reading mechanisms and dig deeper into language comprehension processing in real time. This method became popular in psycholinguistic research, since it was "as similar as possible to normal reading" (Mitchell & Green, 1978, p.610). The basic assumption underlying this technique is that "eyes can be a window on cognition" (Jegerski, 2014, p.23). According to the eye-mind assumption proposed by Just and Carpenter (1980, p.330), the amount of time it takes for the participants to read a word reflects the amount of time that it takes to process the word. Thus, the participants make longer pauses at particular points in a sentence where the processing load is greater. Further research has shown that the relationship between RTs and processing is far more complicated, but overall, the basic assumption still holds, and RTs can be interpreted to draw conclusions about the cognitive processing of language, such that longer RTs are interpreted as processing difficulties, while faster RTs are taken as a sign of ease of processing (Jegerski, 2014; Marsden et al., 2018).

SPR tasks consist of a cue, stimulus, and distractor. The cue is a fixation cross (+) that appears on the screen so that the participants' gaze is fixed on the point where the first word appears. Then the participants see the stimulus, and they can read the next words by pressing a button. In addition to the target stimuli, SPR tasks also consist of noncritical items, commonly referred to as *fillers*, which do not involve experimental manipulation. The fillers appear only in one condition – whether they are all grammatical or ungrammatical – and have the purpose of diverting participants' attention from the target stimuli so that the research objectives are obscured. The distractor is usually a binary-choice question (commonly a Yes/No question) that comes after some or all trials. The purpose of the distractor questions is to give the

participants a goal for reading the stimuli, so that they pay attention to them instead of only pressing the button mechanically. The most common distractor questions are acceptability judgements and meaning-based comprehension questions. Researchers use acceptability judgements, because they provide additional offline data that might be very interesting. However, scholars are concerned about the possible effects of metalinguistic tasks such as acceptability judgement questions on the online processing, since using this type of questions increases the use of explicit knowledge. There is empirical evidence from a study by Leeser, Brandl, and Weissglass (2011) that shows participants' online processing behavior can vary depending on the distractor question; participants were more sensitive to grammatical violations when the distractor was an acceptability judgement, but not when it was a meaning-based comprehension question. Thus, it is recommended to opt for a meaning-based comprehension question as a distractor in SPR tasks (Keating & Jegerski, 2015, p.10-12).

SPR is a general term for various formats of this task. There are three types of presentation in SPR: 1) cumulative presentation, 2) linear non-cumulative presentation, and 3) center non-cumulative presentation (Bloom & Unsworth, 2010). In the cumulative presentation, all regions of the sentence are masked with dashes. Before each sentence starts, a fixation cross appears on the screen. The sentence starts with the first word. By pressing the button, the first word remains, and the second word of the sentence appears on the screen. As the participants read the next words, the previous words remain on the screen, and they can go back and read them. An example of this type of presentation is provided in Figure1 for the sentence *The confident engineer maintained the debate would be easy to win* (taken from Jegerski, 2014, p.46).

+

The ---------- ------ ----- --- ---------- ------- --------- -- ---------- ------.

The confident ------ ----- --- ---------- ------- --------- -- ---------- ------.

The confident engineer -- --- --------- ------- --------- -- ---------- ------.

The confident engineer maintained -- -------- ---------- -- -------- ------.

The confident engineer maintained the -------- ------- -- --------- ------.

The confident engineer maintained the debate -------- -- -------- ------.

The confident engineer maintained the debate would -- --------- ------.

The confident engineer maintained the debate would be -------- ------.

The confident engineer maintained the debate would be easy --- ----.

The confident engineer maintained the debate would be easy to -----.

The confident engineer maintained the debate would be easy to win.

Figure 1 - Cumulative Presentation of SPR with Word-By-Word Segmentation.

In the linear non-cumulative presentation, the words disappear when a new word appears on the screen. This means that the participants cannot go back and read the previous regions. In Figure 2 we can see this type of presentation for the same sentence (taken from Jegesrki, 2014, p.47).

+

The ---------- ------ ----- --- ---------- ---------- ---------- -- ---------- ------.

------ confident ------- ------ --- ---------- ---------- ---------- -- ---------- ------.

----- ---------- engineer-- --- ---------- ---------- ---------- -- ---------- ------.

----- ---------- ------ ----- maintained ---- ---------- ---------- -- ---------- ------.

----- ---------- ------ ----- --- ---------- the --------- -------- -- ---------- ------.

----- ---------- ------ ----- --- ---------- ----- debate -------- -- ---------- ------.

----- ---------- ------ ----- --- ---------- ----- ---------- would -- ---------- ------.

----- ---------- ------ ----- --- ---------- ----- ---------- -------- be -------- ------.

----- ---------- ------ ----- --- ---------- ----- --------- --------- --- easy --- -----.

----- ---------- ------ ----- --- ---------- ----- --------- ---------- --- ----- to -----.

----- ---------- ------ ----- --- ---------- ----- --------- ---------- --- ----- --- win.

Figure 2 - Linear Non-Cumulative Presentation of SPR with Word-By-Word Segmentation.

In the center non-cumulative presentation, the words appear at the center of the screen, as shown in Figure 3 (taken from Jegesrki, 2014, p.48).

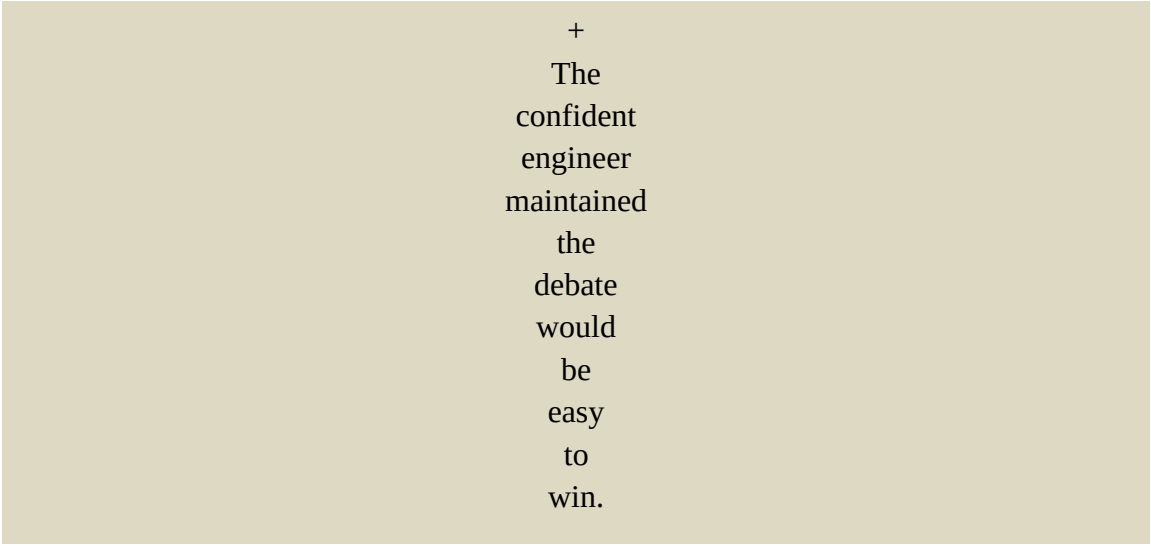


Figure 3 - Center Non-Cumulative Presentation of SPR with Word-By-Word Segmentation.

In the present study, it was decided to use the center non-cumulative display instead of the linear non-cumulative display (which is a more common way of presentation in SPR) after conducting several pilot studies (see section 3.6 for more details about the pilot studies). In the SPR tasks with linear non-cumulative presentation, the participants are able to see the length of each sentence. Furthermore, the cumulative presentation is problematic, because for instance, readers can develop a reading strategy in which they can uncover several regions of a sentence and read them altogether instead of reading the segments one at a time (Jegerski, 2014; Ferreira & Henderson, 1990; Just et al., 1982). Furthermore, the format of segmentation in SPR is of importance. In SPR tasks, word-by-word segmentation is preferred since it provides more precise data. Word-by-word segmentation collects more data at different points per stimulus compared to phrase-by-phrase segmentation, and the data can also be easily converted into phrase-by-phrase format by summing the RTs of the different regions. However, phrase-by-phrase segmentation is more similar to natural reading, but once an experiment has been run in phrase-by-phrase mode, there is no way to break down the reading times into word-by-word reading times unless the experiment is run again. Moreover, the phrase-by-phrase segmentation introduces some complications that can affect processing behavior through the way that

particular words are grouped together and also by the length of the phrases (Jegerski, 2014, p.30-31). Both formats of segmentation are illustrated in [14] and [15].

[14] *Word-by-word segmentation*

I / saw / that / tall / man / in / the / car / over / there.

[15] *Phrase-by-phrase segmentation*

The mother called / her daughter / while / she was baking / an apple cake.

There are several motivations for the choice of SPR in this study. First, online methods of data collection are preferred to offline methods because they provide information about sentence comprehension at different points so that researchers can investigate what happens at a particular point in a sentence at the exact moment when those words or phrases are perceived. Second, online methods measure participants' behavior in real time, meaning they tap into participants' implicit knowledge and do not allow the use of conscious explicitly learned knowledge. Similarly, there are other online methods that investigate online processing such as Event Related Potentials (ERP) and eye-tracking, both of which have time-consuming procedures and high cost, while SPR introduces relative ease of administration and low cost (Keating & Jegerski, 2015; Marsden et al., 2018).

3.2.2 Acceptability Judgment Task

Another task for data collection in this study was an acceptability judgement test (AJT), which is a widely used method in linguistics due to its easy administration (Dabrowska, 2010). In the acceptability judgement tests, participants decide whether sentences are acceptable or unacceptable.

Acceptability judgment tests are also referred to as grammaticality judgement tests, and these two terms are often used interchangeably in the literature. However, the former is used throughout this study, since acceptability and grammaticality are two different concepts. Chomsky (1965, p.11) defines –acceptability as "… a concept that belongs to the study of performance, whereas grammaticalness belongs to the study of competence." Therefore, as Leivada and Westergaard (2020) suggest, there may be sentences that are acceptable, although they are ungrammatical, and sentences that are unacceptable, but they are grammatical. This

suggests that grammaticality is only one of the factors that affects acceptability in addition to other performance-associated factors such as memory limitations and processing constraints (Leivada &Westergaard, 2020, p.2). For instance, sentence [16a] is a grammatically correct sentence, because it does not violate any grammar rules; however, grammaticalness does not guarantee acceptability. Speakers would not judge sentence [16a] as acceptable, but it is a grammatical sentence, in which no grammar rule is violated. Likewise, [3b] may be an ungrammatical sentence, but it is acceptable. Example [16a] is taken from Barton et al. (1987, as cited in Leivada &Westergaard, 2020, p.3) and [16b] is taken from Dabrowska (2010, p.4).

[16] a) Dogs dogs dog dog dogs.

b) Watched some TV, then went to bed.

One of the criticisms against AJTs is that this method of data collection is not reliable, because the participants' judgments may be affected by non-linguistic factors, and therefore it is easy to make Type I or Type II errors (Sprouse & Almeida, 2011). Type I errors, or 'false positives' or 'false alarms', are when the result shows that, for example, two experimental groups are different, but the result is due to a sampling error, and there is no significant difference. Therefore, the null hypothesis is rejected in this case although there is no real difference (Levshina, 2015, p.13). A Type II error, also known as 'false negatives', is when the null hypothesis is accepted while there is in fact a real difference between the groups (Levshina, 2015, p.13). The fact that Type I or Type II errors may occur has been argued to be one of the drawbacks of AJTs.

However, Sprouse and Almeida (2012) provided evidence that shows AJTs are in fact reliable. In their study, they tested all the data points in the textbook by Adger (2003), and the results showed that the maximum rate for the possibility of false positive is only 2%. Thus, it was shown that the AJTs have a minimum replication rate of 98%, and therefore they are reliable.

In the AJT in this study, the participants were asked to judge the acceptability of the sentences on a Likert scale from 1 to 4, in addition to the option 'I don't know', as presented in Figure 4.

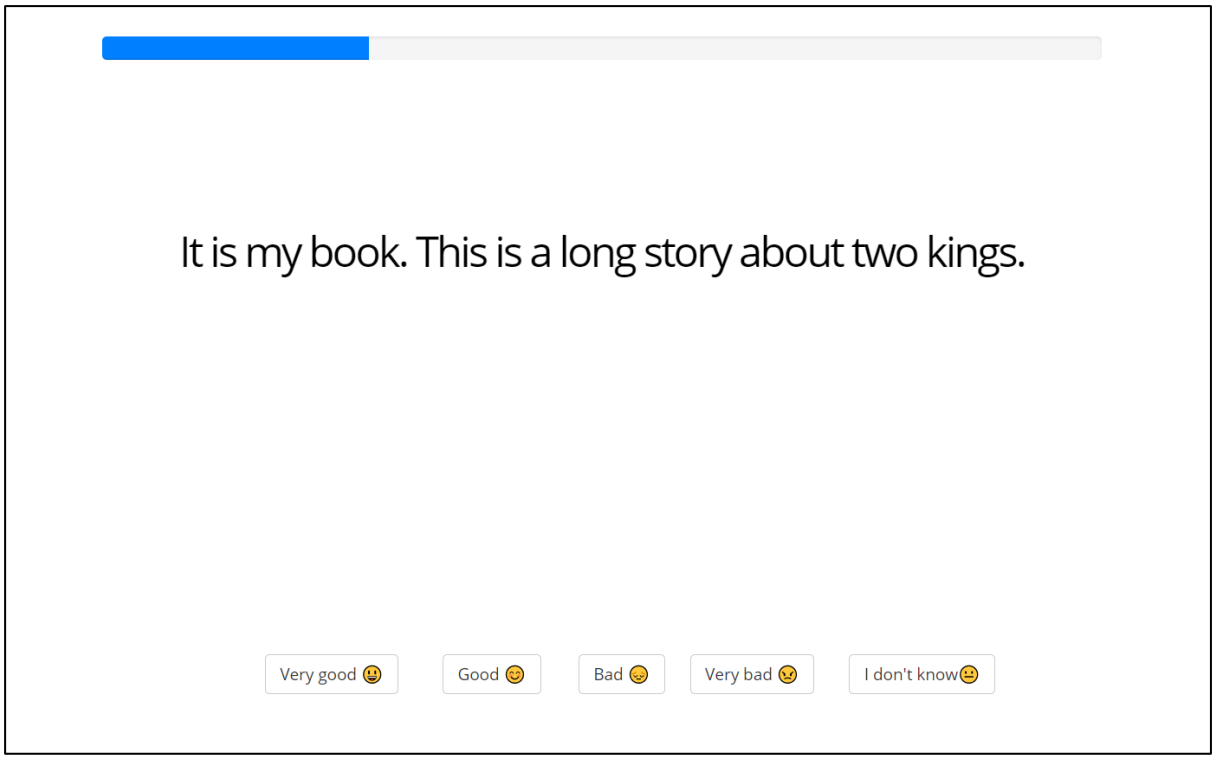


Figure 4 - The Likert Scale in the AJT.

The AJT in this study provides additional data on the participants' comprehension of English. Moreover, this is an opportunity to compare the data from an online task (SPR) with the data from an offline task (AJT).

3.3 Procedure

All parts of the experiment were built and hosted by Gorilla Experiment Builder (www.gorilla.sc), which is a web-based platform that makes recruiting participants from different parts of the world relatively easy (Anwyl-Irvine et al., 2019). The whole experiment was done on-line due to the pandemic situation. An online experiment was the most convenient option also because the participants took part from different cities in Norway and Iran. The experiment took approximately 30-35 minutes to complete.

Depending on the groups, the participants went through different tasks. The L3 group went through all of the steps outlined below, whereas the two control groups did not have the last two audio acceptability judgement tasks.

1. A consent form.
2. A self-paced reading task in English.

3. An English proficiency test.
4. A background questionnaire.
5. An acceptability judgment task in English.
6. A mini audio acceptability judgement task of the properties under investigation in Norwegian.
7. A mini audio acceptability judgement task of the properties under investigation in Persian.

The consent form was integrated in the experiment. It consisted of an information section in which necessary information about the experiment was provided (see Appendix 1) and a consent form in which the participants' parents declared that they agree with their child's participation in this experiment. After completing the consent from, the participants were given the instructions (in either Norwegian or Persian) about the SPR task, and after finishing three practice trials, they entered the main SPR task. All items in the SPR were randomized. The third part of the experiment was a subset of a standardized Oxford proficiency test which originally had 40 items, but it was shortened to 29 items (see Appendix 2). As shown in figure 5, the proficiency test is a multiple-choice task where the participants choose the best answer among the three options provided to complete a sentence. The participants were able to choose only one option for each sentence, and they could not move on to the next task unless they had completed this test.

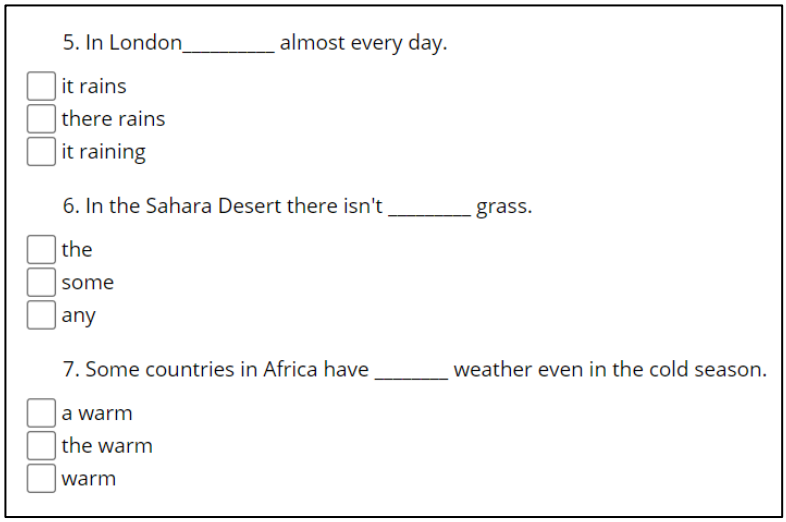


Figure 5 - Multiple-Choice Proficiency Test.

The next task was a short background questionnaire. The participants were asked to answer a few questions about their age, sex and language use with their mother, father, and friends (see Appendix 3). The background questionnaire was either in Norwegian or Persian depending on the group. The participants then moved on to the AJT that contained the same test items as the SPR task. Before going through the main AJT, the subjects were provided with necessary instructions (in either Norwegian or Persian) and also had two practice trials. All sentences were randomized, and the participants saw one sentence at a time, and they were not able to go back and change their responses.

The final parts were two mini audio acceptability judgment tasks, one in Persian and one in Norwegian (see Appendix 5 and 6 for the list of items). Since most of the heritage speakers are unable to read or write in their heritage language, it was decided to test the language properties in their background languages through an audio grammaticality judgment task, where they listened to some sentences and rated their acceptability on a Likert scale. Before starting the task, the participants were given the instructions and were asked to turn up the volume on their laptop or computer. They did not have any practice trials since the tasks were quite simple. The purpose of these tasks was to see if the heritage speakers had acquired the properties in this study in their previously acquired languages.

3.4 Test Items

The SPR task in this study consisted of 60 critical items in two conditions: 30 grammatical and 30 ungrammatical items. Therefore, there were 12 items per property (6 grammatical, 6 ungrammatical) (see Appendix 4 for the complete list). All test items were checked by a native English speaker.

Since the sentences with the gender property required a context sentence, it was decided to add a context sentence to all test items in order to be consistent across all test items. All of the test items consisted of simple, frequently used words, which were taken from a word frequency list (Word Frequency, 2014). It was difficult to keep the same length in all items across all properties. Thus, items within each property have the same length and structure. The length of test items for each property is as follows: gender 13-15 syllables, definiteness 13-15 syllables, Adj-N 14-16 syllables, SV 14-16 syllables, and Adv-V 14-16 syllables. All test items had the same number of regions within each property. The context sentence was presented as

one region (in one chunk, not word by word) in all test items. The critical region (or the region of interest) was never at the beginning of the sentence nor at the end. There were always 2-3 regions before the critical region and also 2-3 regions after, because processing a particular region in a sentence often continues or "spills over" onto the following regions, which is known as the *spillover effect* (Keating & Jegerski, 2015, p.6; Rayner & Duffy, 1986). Table 7 shows examples, divided into different regions, from all properties under investigation, both grammatical and ungrammatical. The critical regions are underlined.

Table 7 - Target Stimuli in Different Properties.

Language Property	Grammatical	Ungrammatical
Gender	My dad works in a hospital/ because/ he/ is/ a/ doctor.	* My uncle doesn't walk home/ because/ she/ has/ a/ car.
Definiteness	Sara has a hat./ Mary/ says/ the/ hat/ is/ very/ big.	*Andy has a phone./ Mia/ says/ phone/ is/ very/ old.
SV	I can hear Tim and John./ Today/ Tim/ laughs/ very/ loudly.	* My family is happy./ Today/ feels/ grandpa/ very/ well.
Adv-V	My friends like running./ Billy/ always/ runs/ very/ fast.	* Andy and Sam can sing./ Andy/ sings/ often/ very/ well.
Adj-N	We need a ruler./ Mia/ has/ a/ long/ ruler/ on/ that/ desk.	* It's cold today outside./ John/ wears/ a/ jacket/ warm/ in/ that/ room.

The items in grammatical and ungrammatical conditions were different sentences in order to avoid repetition effects. When participants read the same item in slightly different versions (e.g., different word order), this could result in an unnatural response because they have seen that sentence before. Furthermore, reading the same sentences in different conditions would increase the likelihood that participants become aware of the target of the experiment.

In addition to the target items, there were 60 grammatical filler sentences without experimental manipulation in the SPR task (see Appendix 5). These were unrelated sentences in which the target language properties were avoided. The reason why the filler sentences were grammatical is to make the ungrammatical target items stand out in the SPR task. The fillers

also start with a context sentence, and they were the same length as the target stimuli. Two examples of the filler sentences can be found in [17].

[17] *Filler sentences*

a) Exams start tomorrow. Students are studying very hard.
b) Ed and Lisa are teachers. Ed teaches math at school.

The same 60 target items were tested in the AJT, without the fillers, since the sentences in different conditions serve as fillers for each other. The last tasks were the two audio AJTs in Norwegian and Persian. The Norwegian task consisted of 10 items: 5 grammatical and 5 ungrammatical. The Persian task consisted of 6 items: 3 grammatical and 3 ungrammatical. The smaller number of items in the Persian task is due to the fact that Persian does not have gender or a definite marker, therefore these two properties could not be tested in Persian. The sentences were recorded by Norwegian and Persian native speakers using the iPhone voice recording application. The audios were then edited by the computer program Audacity to ensure that there was no noise in the background and that the pauses between words were of the same length across all items.

3.5 Participants

Three groups of English learners were tested in this study: 1) 29 Persian-Norwegian heritage speakers who were learning English as an L3 at school in Norway (L3 group), 2) 24 Persian speakers who were learning English as an L2 in Iran (Persian controls), and 3) 29 Norwegian speakers who were learning English as an L2 in Norway (Norwegian controls). The target age group was originally 11-14. However, there were older participants in the Persian control group and younger ones in the Norwegian control group. This was done, because it was very important for the groups to be comparable in terms of English proficiency. Therefore, after adding the older and younger participants to the control groups, a logistic regression model was fit in which accuracy was predicted by group as fixed effect and participants and questions as random effects. The results showed no significant difference between the mean English proficiency of the groups (see Appendix 7 for the regression model).

The Persian-Norwegian bilinguals were recruited from Iranian families from different cities in Norway, and they were asked to do the experiment in their spare time. The L1 Persian

learners were recruited from a private school in Bandar Abbas, Iran, and they were asked to do the experiment in their spare time or at school. The L1 Norwegian learners were recruited from two schools in Tromsø, Norway, and they were asked to do the experiment in their spare time or at school. Table 8 shows the participants' details in each group.

Table 8 - Description of the participants.

Groups	n	Country of residence	Mean Age (in years)	AoO[4] in English (M in years)	LOC[5] with mother/father	LOC with friends	Mean English Proficiency
L3	29	Norway	12.2	5.2	Per[1]/Per (n =17) Nor[2]/Nor (n =1) Per/Per&Nor (n =1) Nor&Per/Nor&Per (n =9)	Nor (n =20) Eng[3] (n =4) Nor&Eng (n =5)	0.71
Persian Control	24	Iran	15.7	7.8	Per/Per (n =24)	Per (n =24)	0.66
Norwegian Control	29	Norway	12.4	5.8	Nor/Nor (n = 28) Nor/English (n = 1)	Nor (n = 21) Nor&Eng (n = 8)	0.72

[1]Per = Persian, [2]Nor = Norwegian, [3]Eng = English, [4]AoO = Age of Onset, [5]LOC = Language of Communication

3.6 Piloting

Before the main experiment, several pilot studies were conducted. In this section, all pilot studies and their results are discussed. Several modifications were made based on the results from each pilot, which will be explained in detail through the rest of this section.

The RTs were trimmed before visualization in all pilot studies. According to Jegerski (2014, p.39-40), the lower cutoff is 200 ms, since any RT less than that typically reflects unintentional button press. The higher cutoff is 3000 ms for native readers and 4000 ms for non-native readers. Therefore, any RT less than 200 ms and more than 3000 ms was removed from the dataset (the higher cutoff for the third pilot study was 4000 ms, since the participants were non-native readers) and mean RTs were visualized in all pilot studies.

The first pilot study included a self-paced reading task with distractor questions and a linear non-cumulative display and a short background questionnaire about the participants' age, sex, and language of communication with their parents and friends. The participants were 6 adult native English speakers at UiT the Arctic University of Norway. The participants had only

English as their L1, and they all used English as the language of communication with their mother, father, and friends. The purpose behind carrying out the pilot study with native English speakers was mainly to investigate if they made a distinction between the grammatical and ungrammatical trials. The SPR task included meaning-based questions as distractors after some of the sentences (a total of 12 questions). In the SPR tasks, it is generally expected that there would be longer reading times in (or after) the critical regions for the ungrammatical condition than for the grammatical condition. The results of the first pilot were more or less in line with the expected general trend; however, the difference between the mean RT of the grammatical and ungrammatical trials in the critical region was not significant in most properties (see Appendix 8 for the graphs). It was difficult to tell whether the results were due to the participants' lack of motivation (since the participants did not receive any rewards) or a problem with the design of the test items. Therefore, a second pilot study was conducted.

For the second pilot study, everything was kept the same, except for the method of recruiting participants. Prolific (https://prolific.co/), an online recruitment platform, was chosen as the new method of recruitment. Recruiting participants via Prolific was more efficient for two reasons: 1) through Prolific it was possible to pay all participants for the duration of time they spent on the experiment (something which was not done in the first pilot), and 2) the recruitment was done quite fast. Again, the participants were 6 adult native English speakers who had only English as L1 and the only language of communication. The results of the second pilot showed an opposite trend than what was expected. The RT of the ungrammatical items was faster than grammatical ones in the critical region in three out of five properties (see Appendix 8). Since the results could not be due to a lack of incentive, another pilot was conducted with a different population in order to find out if the test items were the problem.

The design of the third pilot study was the same as the first and second one, only the participants were different in order to see if they behave differently than the native English speakers. Participants in this pilot study were 5 adult native Persian speakers in Iran. They all had Persian as their L1, English as L2. The results of the third pilot study were quite different from the first and second pilot. For gender and Adv-V properties, there was a significant difference between the RTs of the grammatical and ungrammatical items in and after the critical region (see Appendix 8). For the Adj-N word order the RT of the ungrammatical items was significantly longer than grammatical items in the critical region only. For the definiteness

property, the RT of the ungrammatical items was slightly longer than grammatical items in and after the critical region. For the SV property, the RTs of both grammatical and ungrammatical items were quite the same in the critical region, however, a longer RT for ungrammatical than grammatical items in the spillover regions was recorded. Therefore, it was concluded that the issue is not the test items. Since the three pilots showed the problem is not the participants' lack of incentive or a flaw in the test items, it was decided to change the design of the SPR task for the next pilot study to find out if that makes any difference in the result.

In the fourth pilot study, the participants were 5 native English speakers who had only English as L1 and the only language of communication. There were two main changes in this pilot study: 1) different presentation in the SPR task, and 2) the addition of filler sentences. It was decided to change the SPR display from linear non-cumulative to center non-cumulative in order make the task more challenging for the participants. As mentioned before (see section 3.2.1), in the center non-cumulative display (unlike the linear non-cumulative) the readers cannot see the length of sentences or words, so they do not know what to expect next, and that makes the task more challenging. Furthermore, the previous pilot studies did not include filler sentences because the test items from each property served as fillers. However, it was decided to add 60 grammatical filler sentences in addition to the 60 target test items to make the ungrammatical items stand out. The results were quite the same across all properties (see Appendix 8). The RT of the grammatical and ungrammatical items was almost the same in the critical region and the regions after in most properties, and the RT of the ungrammatical items was even slightly faster than the grammatical items for the SV property, which was not expected. After conducting four pilot studies, it became clear that obtaining significant result from the SPR task was unlikely, for several reasons such as participants' lack of focus on the task. Therefore, it was decided to add an acceptability judgement task to the next pilot study in addition to the SPR task.

The participants in the fifth pilot study were 5 adult native English speakers. I also asked my cousin who is 13 years old to participate in Iran. The native English speakers had only English as L1 and they all used English as the language of communication with their mother, father, and friends. My cousin had Persian as the only L1 and English as L2 and her only language of communication was Persian. The reason why I asked my cousin to participate was to see if the instructions were easy to understand and follow for the target age group.

After the native English speakers completed the experiment, the data were visualized in graphs for the SPR task and the AJT (see Appendix 8). The results showed that the reading time for grammatical and ungrammatical items for the SV property was the same in the critical region, with a slight increase for the ungrammatical condition in the last region. The reading times for the Adv-V property were also fairly equal, with a very slight increase in reading time for the ungrammatical condition in region 5. For the Adj-N property, the opposite trend than what was expected occurred. The reading time in the critical region for the grammatical condition was slightly longer than for the ungrammatical trials, but the reading time for the ungrammatical condition increased in the last two regions. For gender and definiteness, the reading time for ungrammatical trials was longer than the grammatical trials in the critical region and even in the spillover regions, which was originally expected. The AJT results showed that participants discriminated between grammatical and ungrammatical sentences, as the mean score for acceptability of grammatical sentences was significantly higher than that of ungrammatical sentences for all properties (see Appendix 8). Minor modifications were made based on feedback from my cousin and other participants. My cousin and another participant thought that their voice was being recorded during the SPR task. Therefore, a part was added to the SPR task instruction to reassure participants that their voice would not be recorded. They had no other difficulties with the instructions or the design of the experiment.

In sum, it was concluded that the SPR results were influenced to a large extent by the participants' lack of focus, or the results were due to a too low number of participants. Additionally, according to Kaan et al., (2019) and Fine et al., (2013), who tested English native speakers in their native language via SPR tasks, the native speakers rapidly adapted to the unexpected structures, and therefore showed a decrease in processing difficulty. Therefore, this could also be the case with the English native speakers in the pilot studies of the present study; they adapted to the ungrammaticality, and that is why the difference in RTs between the grammatical and ungrammatical conditions did not reach significant. However, since the Persian L2 learners of English did differentiate significantly between the grammatical and ungrammatical sentences in some of the properties, the participants in the main study may distinguish between the grammatical and ungrammatical conditions. Moreover, the fact that the AJT was included in the experiment in addition to the SPR task makes the design of the study less risky, as participants may get disturbed or distracted by noise or other things going on in the environment during the SPR task, which could subsequently lead to longer or faster reading

times. Therefore, before starting the SPR task, it was emphasized that participants should sit in a quiet environment where they would not be disturbed by anything.

3.7 Chapter Summary

This chapter first presented the objectives of the present study, which were twofold: the source of CLI, and whether CLI occurs on a property-by-property basis or wholesale. Predictions were then made based on the two dominant models in the field of L3A, i.e., TPM and LPM.

While the TPM argues that the L3 participants would be influenced exclusively by Norwegian due to the typological similarity between Norwegian and English, and that the parser will copy the whole language representation (in this case Norwegian), the LPM emphasizes that both Norwegian and Persian would influence English due to the structural similarity between the target language (English) and the background languages (Persian and Norwegian) and therefore CLI occurs property by property.

Based on the results of the several pilot studies, major changes in the design of the experiment were made including: 1) the change with the SPR task display from linear non-cumulative to center non-cumulative, 2) addition of the filler sentences to the SPR task, 3) addition of the AJT, and 4) minor changes with the instructions to make them clearer. Applying these changes improved the design of the experiment and made it less risky. The results of this study with an in-depth statistical analysis are discussed in the following chapter.

4 Results

In this chapter, the results of the tasks employed in this study (the SPR and AJT, as well as the mini AJTs) are presented. The data were retrieved from Gorilla and then analyzed using the statistics program R (R Core Team, 2020). The statistical analyses were done using the *lme4* (Bates et al., 2015) and *emmeans* packages (Length et al., 2020). The main focus of this chapter is: 1) to see if there is a significant difference between the RTs of the grammatical and ungrammatical sentences in each property in the SPR task, and if groups are significantly different from each other, and 2) to see if the accuracy rates of the three groups differ significantly from each other in each property in the AJT.

4.1 The Mini AJTs in Norwegian and Persian

The mini AJTs were two short tasks in which the participants in the L3 group were tested in their background languages (Norwegian and Persian), in order to see if they have acquired the properties under investigation in their background languages. Figure 6 shows the results of the mini AJTs. Note that Persian does not mark definiteness and gender, therefore these two properties were not investigated in Persian.

As evident from Figure 6, the accuracy scores of the L3 group in the Norwegian test were above the set acquisition threshold (>70%) in all properties, except for the definiteness property, where they scored 58%. Similar to the Norwegian test, the L3 group were highly accurate in Persian too, since their accuracy scores were all above the set acquisition threshold (>70%) in the Persian test as well. Thus, the results of the mini AJTs indicated that the L3 group successfully acquired four out of five properties under investigation since the accuracy scores of four of the properties were above 70%. However, they were not as accurate in the definiteness property, which suggests this property had not been fully acquired yet. Overall, the results of the mini AJTs showed that the L3 learners clearly distinguish between Persian and Norwegian.

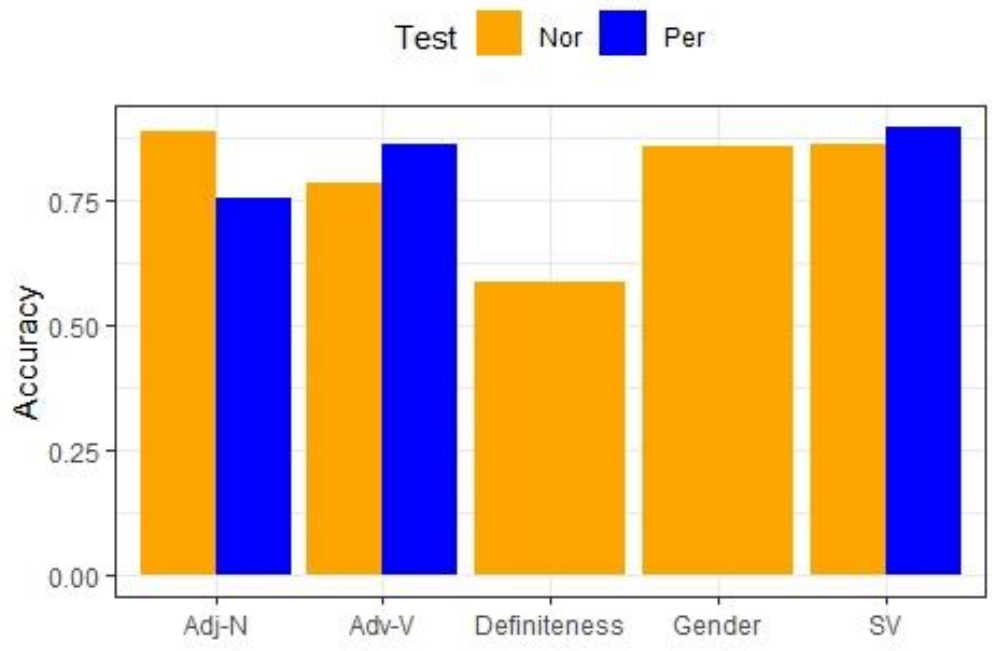


Figure 6 - Accuracy Rates in the Mini-AJTs in the L3 group.

4.2 The Self-Paced Reading Task

Before analyzing the reading times (RTs), data trimming was done. Data trimming is the process of cleaning data for outliers. In order to identify the outliers, the standard deviation approach was used, in which RTs that fall greater than two standard deviations away from the mean were judged as outliers and then removed from the dataset (Jegerski, 2014).

First, RTs for the critical regions were averaged in each property per group. Then, the RTs of the critical regions and the immediately following regions were summed and illustrated for each property per group. Summing the critical regions and the immediately following regions was done in order to avoid misleading results in the properties where word order was investigated. To my knowledge, word order has not been investigated through SPR and therefore this approach was used in order to avoid misleading results. I explain the complication below.

Consider the following examples in a SPR task:

– I /have /a /red /pen /on /my /desk.
R4 R5 R6 R7

– I /have /a /hat /blue /in /my /room.
R4 R5 R6 R7

In the analysis, the RT of R4 in the grammatical sentence should be compared to R5 in the ungrammatical sentence. That will leave 3 spillover regions in the grammatical sentence and 2 spillover regions in the ungrammatical sentence, excluding the last region. The last region cannot be a spillover region since it is at the end of the sentence where readers wrap up the sentence and typically RTs have a sharp increase (Jegerski, 2014). Therefore, the RTs were analyzed in two steps: 1) analysis of the RTs only in the critical region (i.e., comparing R4 in the grammatical sentence to R5 in the ungrammatical sentence), 2) analysis of the sum RTs of the critical regions and the immediately following regions in the grammatical and ungrammatical conditions (i.e., R4 + R5 + R6 + R7). This approach shows if the RTs are significantly different in the critical region, and if the RTs of the grammatical and ungrammatical chunk (red pen vs. hat blue) plus the immediately following regions show any effect. Table 9 provides an overview of the critical regions in each condition per property and also the regions which were summed together.

Table 9 - Overview of the Critical and Spillover Regions.

Property	Critical Regions		Critical + Spillover Regions
	Grammatical	Ungrammatical	
Adj-N	R5	R6	R5+R6+R7
Definiteness	R5	R4	R4+R5+R6
SV	R4	R3	R3+R4+R5
Adv-V	R3	R4	R3+R4+R5
Gender	R3	R3	R3+R4+R5

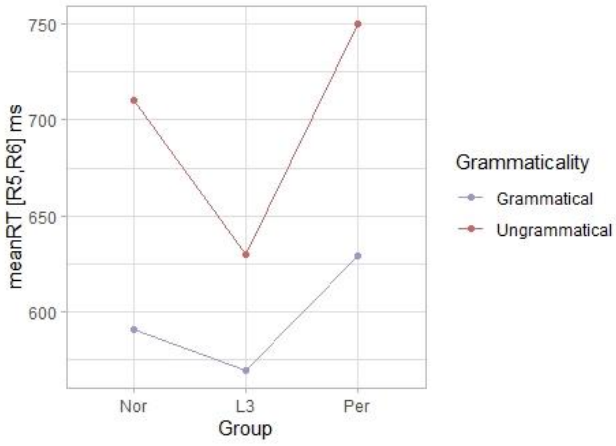
All analyses were done on log transformed RTs. The RTs were analyzed separately for each property. A linear mixed-effect logistic regression model was fit for all properties in which RTs were predicted by group, grammaticality, and their interaction. English proficiency and region length were added as separate fixed effects. Additionally, participants and sentences were added as random slopes.

4.2.1 The Results of the Adj-N, Definiteness and Gender Properties

English patterns with Norwegian in three out of five properties in the present study including: Adj-N, definiteness, and gender, where Persian is different. The mentioned

properties are argued to be problematic for the Persian L2 learners of English (see section 2.4). Therefore, the L3 group and the Norwegian controls are expected to have a significant difference between the RTs of the grammatical and ungrammatical sentences, while this difference is not expected to be significant in the Persian control group based on the TPM on one hand. On the other hand, it is expected that there will be a significant difference between the RTs of the grammatical and ungrammatical sentences in the Norwegian control group, while the Persian controls are not expected to make a significant differentiation based on the LPM. Regarding the L3 group, the RTs are expected to be in the middle of the two control groups in both grammatical and ungrammatical condition.

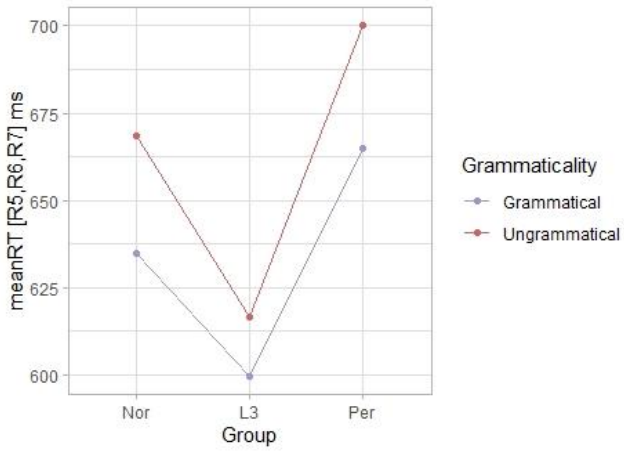
Figures 7 and 8 show the RTs in the critical regions (R5 and R6) and the sum RTs (R5 + R6 + R7) respectively for the Adj-N property. In both figures, the RTs of the ungrammatical condition were slower than the grammatical condition in all groups for the Adj-N property. In the grammatical condition, Persian controls were slower compared to the other two groups, while the L3 group was faster and Norwegian controls were in the middle. The same trend as in the grammatical condition was observed in the ungrammatical condition, however, the groups were statistically the same in both conditions and figures.



– We need a ruler./ Mia/ has/ a / long/ ruler/ on/ that/ desk.
R1 R2 R3 R4 **R5** R6 R7 R8 R9

– It's cold today outside./ John/ wears/ a / jacket/ warm/ in / that/ room.
R1 R2 R3 R4 R5 **R6** R7 R8 R9

Figure 7 - RTs in Target Regions for Adj-N.

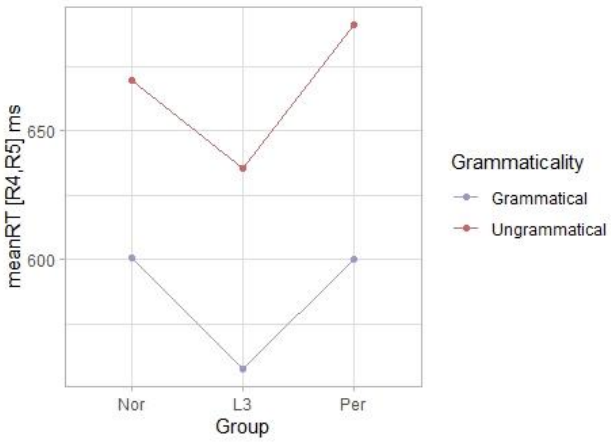


– We need a ruler./ Mia/ has/ a / long/ ruler/ on/ that/ desk.
R1 R2 R3 R4 **R5** **R6** **R7** R8 R9

– It's cold today outside./ John/ wears/ a / jacket/ warm/ in / that/ room.
R1 R2 R3 R4 **R5** **R6** **R7** R8 R9

Figure 8 - Sum RTs for Adj-N.

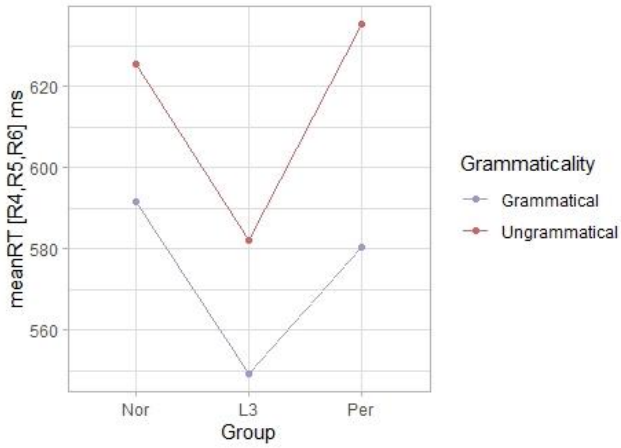
Figures 9 and 10 show the RTs in the target regions (R4 and R5) and the sum RTs (R4 + R5 +R6) respectively for the definiteness property. The RTs of the ungrammatical condition were longer than the grammatical condition in general in both figures. The two control groups had relatively similar RTs in the grammatical condition in the critical region (Figure 4), while the sum RTs of the Norwegian controls were longer than the Persian controls in Figure 5, and the L3 group was faster than both control groups in both figures. In the ungrammatical condition, the Persian controls had a longer RT than the L3 group, and the Norwegian controls were in the middle in both figures. All three groups were statistically the same in the grammatical and ungrammatical conditions in both Figures.



– Sara has a hat./ Mary/ says/ the/ hat/ is/ very/ big.
R1 R2 R3 R4 **R5** R6 R7 R8

– *Andy has a phone./ Mia/ says/ phone/ is/ very/ old.
R1 R2 R3 **R4** R5 R6 R7

Figure 9 - RTs in Target Regions for Definiteness.



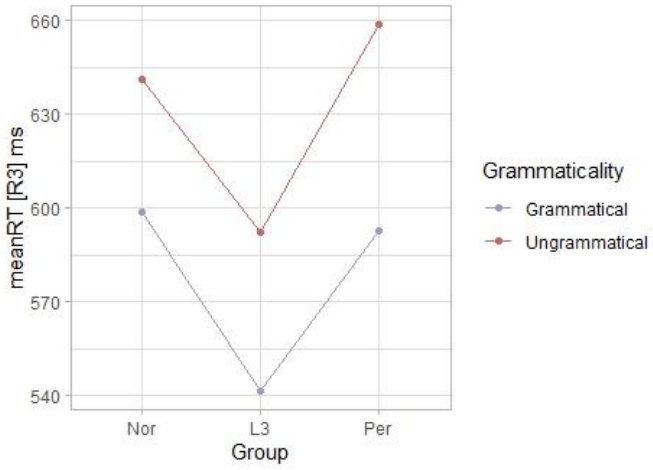
– Sara has a hat./ Mary/ says/ the/ hat/ is/ very/ big.
R1 R2 R3 R4 **R5** **R6** **R7** R8

– *Andy has a phone./ Mia/ says/ phone/ is/ very/ old.
R1 R2 R3 **R4** **R5** **R6** R7

Figure 10 - Sum RTs for Definiteness

The RTs for the target region (R3) and the sum RTs (R3 + R4 + R5) for the gender property are illustrated in Figure 11 and 12 respectively. Generally, the RTs in the ungrammatical condition were longer than in the grammatical condition in all groups in both figures. In the grammatical condition, the Norwegian control group had a longer RT than the L3 group, and the Persian control group was in the middle in the critical region. Figure 12 shows that the Persian controls had a longer RT than the L3 group, and the Norwegian controls were in the middle in the grammatical condition. In the ungrammatical condition, the Persian control group was slower than the L3 group, and the Norwegian control group was in the middle

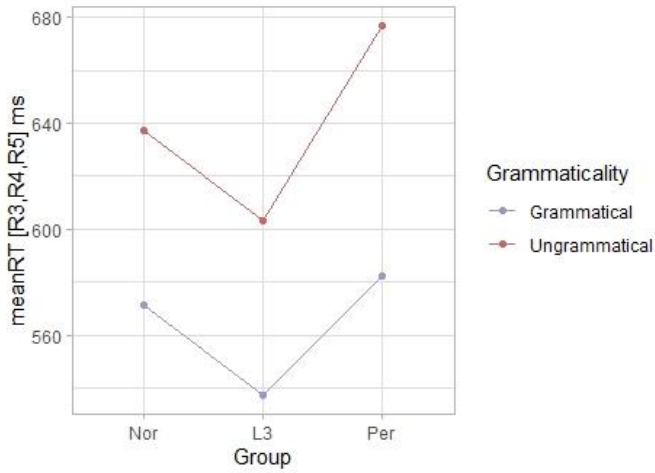
in both figures. There were no statistically significant differences between the groups in the grammatical and ungrammatical conditions in any of the figures.



– My dad works in a hospital/ because/ he / is / a / doctor.
R1 R2 **R3** R4 R5 R6

– My uncle doesn't walk home/ because/ she/ has/ a / car.
R1 R2 **R3** R4 R5 R6

Figure 11- RTs in Target Regions for Gender



– My dad works in a hospital/ because/ he / is / a / doctor.
R1 R2 **R3 R4 R5** R6

– My uncle doesn't walk home/ because/ she/ has/ a / car.
R1 R2 **R3 R4 R5** R6

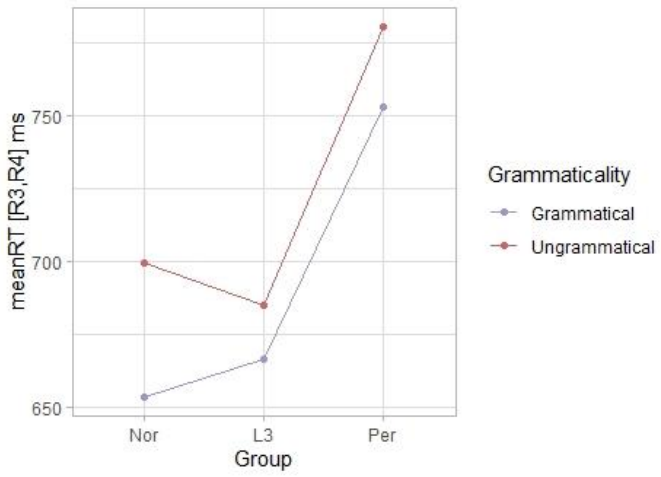
Figure 12- Sum RTs for Gender

4.2.2 The Results of the SV and Adv-V Properties

English patterns with Persian in SV (subject-verb word order in non-subject-initial declaratives) and Adv-V properties. The mentioned properties are argued to be problematic for the Norwegian L2 learners of English (see section 2.4). Therefore, the L3 group and the Persian controls are expected to have a significant difference between the RTs of the grammatical and ungrammatical sentences, while this difference is not expected to reach significance in the Norwegian control group based on the TPM on one hand. On the other hand, it is expected that there will be a significant difference between the RTs of the grammatical and ungrammatical sentences in the Persian control group, while the Norwegian controls are not expected to make a significant differentiation based on the LPM. Regarding the L3 group, the RTs are expected to be in the middle of the two control groups in both grammatical and ungrammatical condition.

Figures 13 and 14 show the RTs of the target regions (R3 and R4) and the sum RTs (R3 + R4 + R5) respectively in the grammatical and ungrammatical conditions for the SV property. The RTs in the ungrammatical condition were longer than the RTs in the grammatical condition in all groups in both figures. As shown in Figure 13, the Persian controls had a longer RT than the Norwegian controls, and the L3 group was in the middle in the grammatical condition,

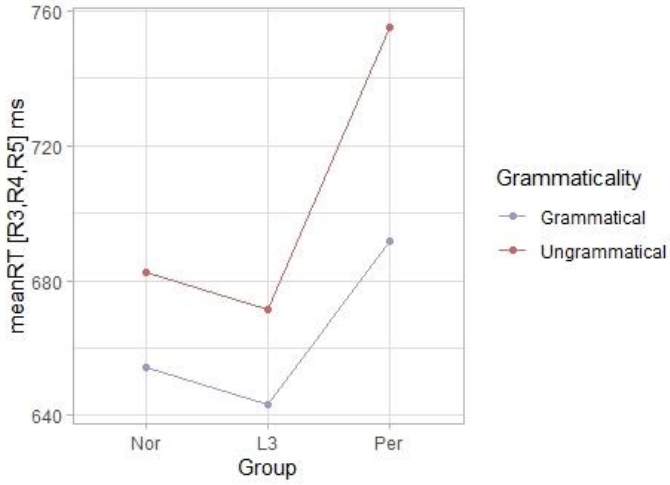
however, the sum RTs of the Norwegian control group was longer than the L3 group in the grammatical condition in Figure 14, and the Persian controls again had a longer RT than the other two groups. In the ungrammatical condition, the Persian controls had a longer RT than the L3 group, and the Norwegian controls were in the middle in both figures. Note that the groups were only numerically different in the grammatical and ungrammatical conditions in both figures, and no significant differences were found.



– I can hear Tim and John./ Today/ Tim/ laughs/ very/ loudly.
R1 R2 R3 **R4** R5 R6

– My family is happy./ Today/ walks/ grandpa/ very/ well.
R1 R2 **R3** R4 R5 R6

Figure 13- RTs in Target Regions for SV

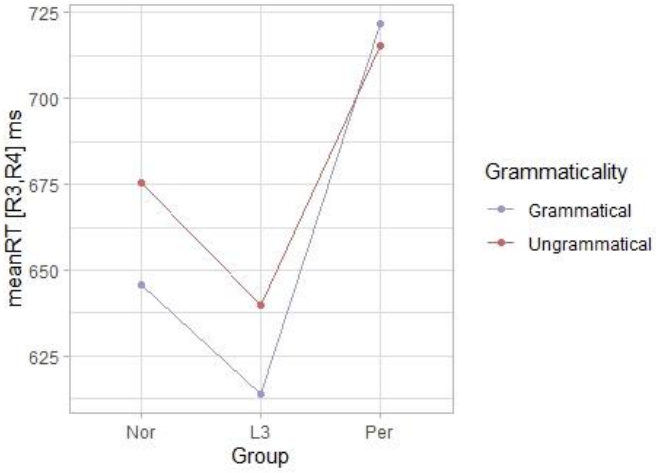


– I can hear Tim and John./ Today/ Tim/ laughs/ very/ loudly.
R1 R2 **R3** **R4** **R5** R6

– My family is happy./ Today/ walks/ grandpa/ very/ well.
R1 R2 **R3** **R4** **R5** R6

Figure 14- Sum RTs for SV

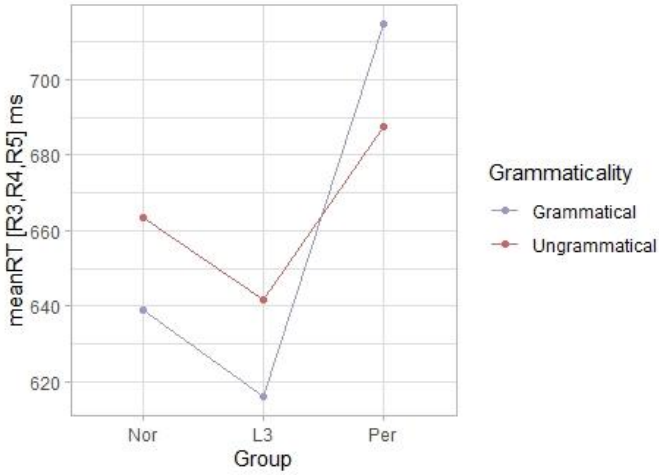
The RTs in the target regions (R3 and R4) and the sum RTs (R3 + R4 + R5) for the Adv-V property are illustrated in Figures 15 and 16 respectively. Figures 15 and 16 illustrate that the ungrammatical RTs were longer than the grammatical RTs only in the L3 group and the Norwegian control group. In the Persian control group, the RT in the grammatical condition was longer than the ungrammatical RT which is not expected. In the grammatical condition, the Persian control group had a longer RT than the L3 group, and the Norwegian control group was in the middle in both figures. The same trend as in the grammatical condition was observed in the ungrammatical condition in both figures. Furthermore, the groups were statistically the same in the grammatical and ungrammatical conditions in both figures.



– My friends like running./ Billy/ <u>always</u>/ runs/ very/ fast.
R1 R2 **R3** R4 R5 R6

– Andy and Sam can sing./ Andy/ sings/ <u>often</u>/ very/ well.
R1 R2 R3 **R4** R5 R6

Figure 15- RTs in Target Regions for Adv-V



– My friends like running./ Billy/ <u>always/ runs/ very</u>/ fast.
R1 R2 **R3** **R4** **R5** R6

– Andy and Sam can sing./ Andy/ <u>sings/ often/ very</u>/ well.
R1 R2 **R3** **R4** **R5** R6

Figure 16- Sum RTs for Adv-V

4.2.3 The Statistical Analysis

The regression models[7] of the critical regions and the immediately following regions only show a significant effect of English proficiency (p-value<0.05) and, in a few cases, region length (p-value<0.05). Therefore, another regression model was fit in order to predict RTs by English proficiency and group and their interaction in each property. Since all the regressions had the same result, it is repeated only once here.

The regression model of English proficiency and group shows a very significant effect of English proficiency (p-value<0.001) and group (p-value<0.05) and an interaction between these two variables. As the English proficiency increased, RTs became significantly faster in the L3 group. Although there was a significant effect of English proficiency in the Persian and Norwegian control groups as well, it was not as significant as in the L3 group. Therefore, higher English proficiency did not necessarily lead to shorter RTs in the control groups as shown in Figure 17. Region length also affected the RTs in a way that the longer the region was, the longer the RT was.

[7] See Appendix 9 for all regression tables of the SPR task.

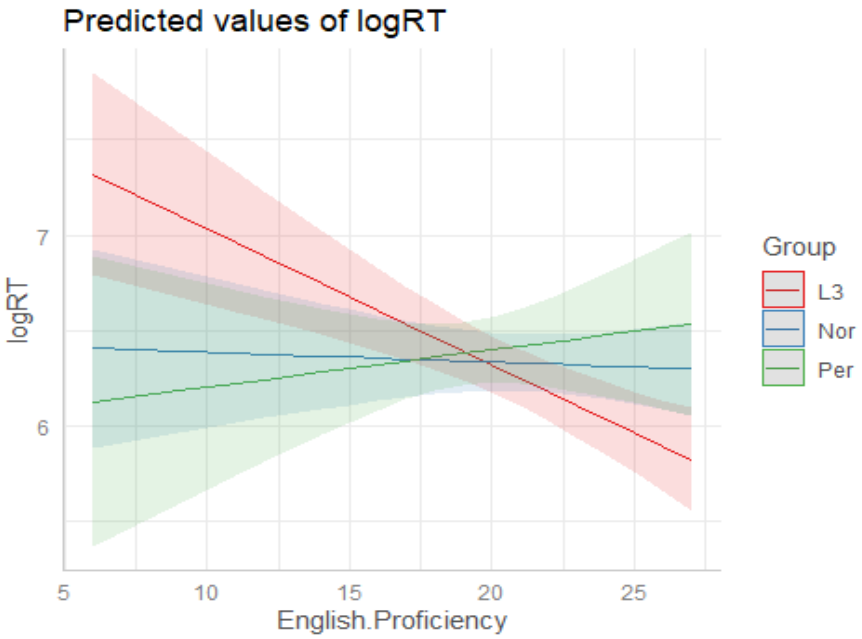


Figure 17- Predicted Values of RTs by English Proficiency

In sum, the results of the SPR task showed that all three groups did differentiate between the grammatical and ungrammatical sentences numerically in all properties, however, the difference did not reach significance, and there were no significant differences between the groups either, since the regression models did not find any significant effect of group or grammaticality or any interaction between these two variables. However, English proficiency and region length were two significant predictors in almost all regression models.

4.3 The Acceptability Judgement Task

Figure 18 shows the mean accuracy of each group in the properties under investigation in the AJT in L3 English. As mentioned in section 3.2.2, the participants rated the acceptability of the sentences on the Likert scale from 1-4, where 1 means very bad and 4 means very good. In the analysis, the scores were treated as a binary variable where 1 and 2 on the Likert scale were equal to 0 (unacceptable) and 3 and 4 were equal to 1 (acceptable). Therefore, for grammatical sentences, an accuracy score of 1 indicates a correct judgement, while 0 means an incorrect judgement. For the ungrammatical sentences, on the other hand, an accuracy score of 0 indicates a correct judgement, while 1 means an incorrect judgement. The data were analyzed

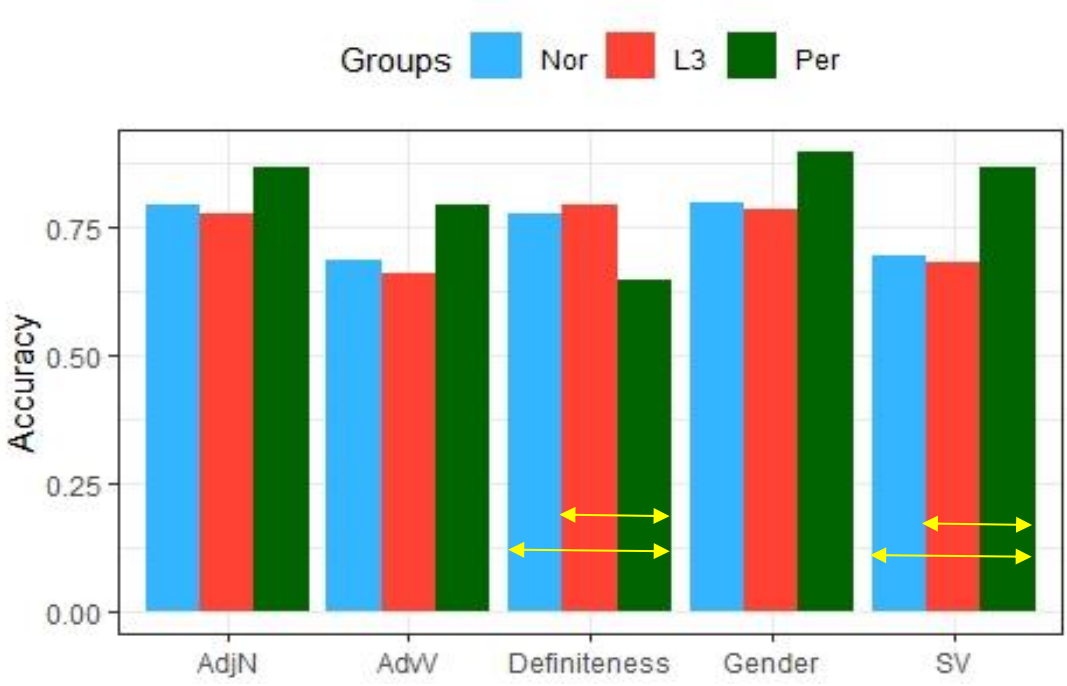


Figure 18 - Accuracy Rates in the AJT of L3 English (significant contrasts between the groups marked with arrows)

using a binomial generalized linear mixed effects logistic regression model[8] to predict accuracy by group, syntactic condition, and their interaction. Age was added as a separate fixed effect. Participants and sentences were added as random slopes.

4.3.1 The Results of the Adj-N, Definiteness and Gender Properties

Norwegian patterns with English in Adj-N, definiteness, and gender properties, where Persian is different, and therefore it is expected that the Persian controls will experience difficulties in the mentioned properties, in a way that it is expected that the L3 group and Norwegian controls will score the highest (the difference between the two groups will not be significant) and the Persian control group the lowest based on the TPM on one hand. One the other hand, it is expected that the Norwegian control group will score the highest and the Persian control group the lowest, while the L3 group is expected to be in the middle based on the LPM.

All three groups were highly accurate in the Adj-N property since their mean accuracy is higher than the set acquisition threshold, which is 70% (Persian controls = 86%, Norwegian controls = 70%, and L3 group = 77%). Although the accuracy rates of the groups are

[8] See Appendix 10 for all regression tables and post-hoc pairwise comparisons of the AJT.

numerically different, the post-hoc pairwise comparison of the model showed no statistically significant difference between the groups.

In the definiteness property, the L3 group scored significantly higher than the Persian control group (79% vs. 64%, p-value = 0.0005), and the Norwegian control group scored 77% in the definiteness property, which is numerically less than the accuracy of the L3 group but not statistically different as predicted by the TPM. The difference between the Norwegian and Persian control groups was also significant (p-value = 0.002).

Figure 18 shows further that all groups were highly accurate in the gender property, where the Persian controls had an accuracy score of 89%, the Norwegian group 79%, and the L3 group 78%. No statistically significant differences were found between the groups.

4.3.2 The Results of the SV and Adv-V Properties

Persian patterns with English in SV and Adv-V properties, where Norwegian is different, and therefore it is expected that the Norwegian controls will experience difficulties in the mentioned properties, in a way that it is expected that the L3 group and Persian controls will score the highest (the difference between the two groups will not be significant) and the Norwegian controls the lowest based on the TPM on one hand. One the other hand, it is expected that the Persian control group will score the highest and the Norwegian control group the lowest, while the L3 group is expected to be in the middle based on the LPM.

In the SV property, the Persian controls significantly outperformed the Norwegian control and L3 groups (86% vs. 69% vs. 67% respectively, p-value = 0.006) as predicted by the TPM. The difference between the Norwegian control group and the L3 group was numerically but not statistically different in the mentioned property.

Similar to the Adj-N and gender properties, all groups had high accuracy scores in the Adv-V property (Persian controls = 79%, Norwegian controls = 70%, and L3 group = 69%). Again, all groups were numerically but not statistically different from each other.

4.3.3 The Statistical Analysis

The regression model showed a significant effect of age (p-value = 0.008) and an interaction between syntactic condition and group. Since there was a significant effect of age in the analysis of the accuracy scores, another regression model was fit in order to predict

accuracy by age and group and their interaction. The results of the regression showed a significant effect of age (p-value<0.05) and group (p-value = 0.004). As shown in Figure 19, younger participants in the Norwegian control group were more accurate than the participants with the same age in the other two groups, and the Persian controls were slightly more accurate than the L3 group in younger age groups across all properties. As age increased, so did the accuracy in the L3 and the Persian control groups. However, the opposite trend was observed in the Norwegian group; as age increased, the accuracy decreased.

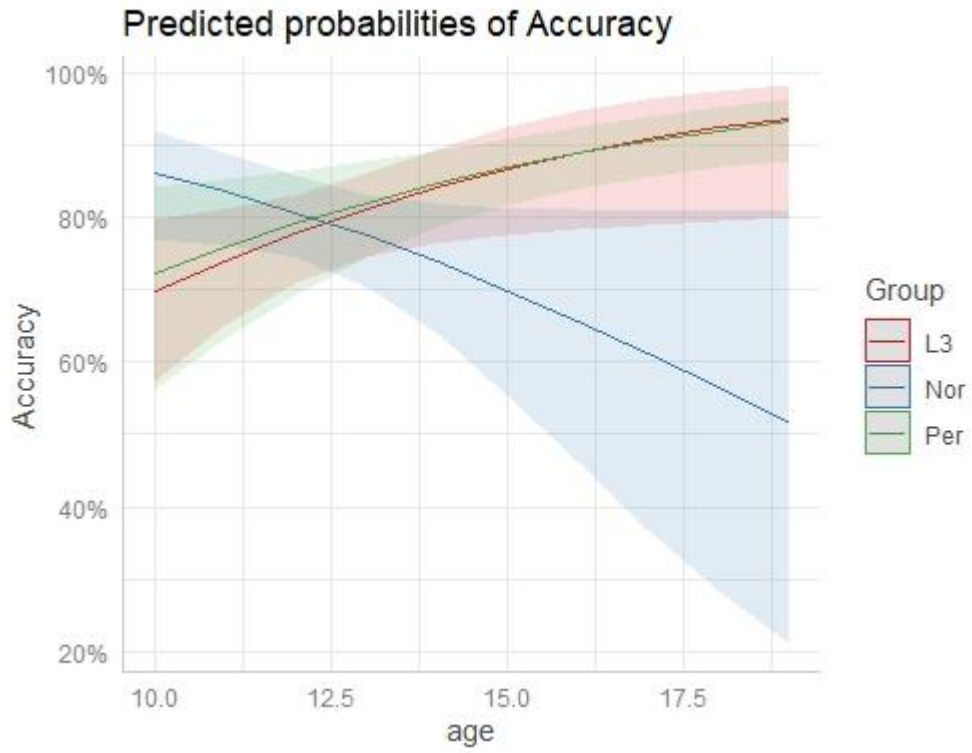


Figure 19 - Predicted Probabilities of Accuracy by Age and Group Across All Properties.

Figure 20 provides an overview of how each participant in different age groups in the Norwegian control group is behaving. It is evident from Figure 15 that the 13 and 14-year-old participants in the Norwegian group scored lower than the younger participants in general. Additionally, there were a few 13 and 14-year-old participants in the Norwegian control group who scored very low, and this probably caused the decrease in the accuracy scores as age increased in Figure 19.

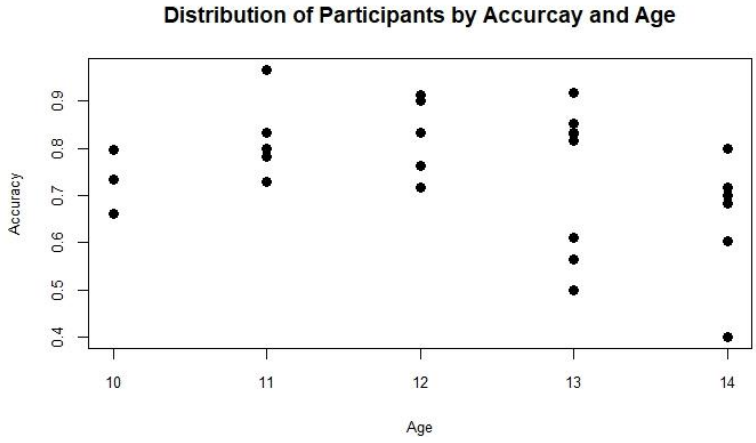


Figure 20 - Distribution of Participants by Accuracy and Age in the Norwegian Control Group.

Table 10 shows the mean English proficiency scores of the different age groups in the Norwegian control group. It is evident from Table 10 that the 13- and 14-year-olds, generally, had higher English proficiency than the younger age groups. Therefore, the low accuracy of the older age groups in the Norwegian control group in the AJT could not be due to low English proficiency.

Table 10 - Mean English Proficiency Scores of the Norwegian Group.

Age Groups	10	11	12	13	14
Mean Proficiency	19.3	20.4	20.3	20.8	21.2

In order to see how English proficiency influences accuracy in different groups, another binomial generalized linear mixed effect logistic regression was fit in which accuracy was predicted by English proficiency, group, and their interaction. The results showed that the English proficiency significantly affects accuracy (p-value<0.001) . As shown in Figure 21, as the English proficiency increased, so did the accuracy in all three groups. It can be concluded from Figure 21 that English proficiency influenced accuracy in the Persian control group more

than the other two groups, because as the English proficiency increased, the Persian controls had higher accuracy than the L3 and Norwegian control group.

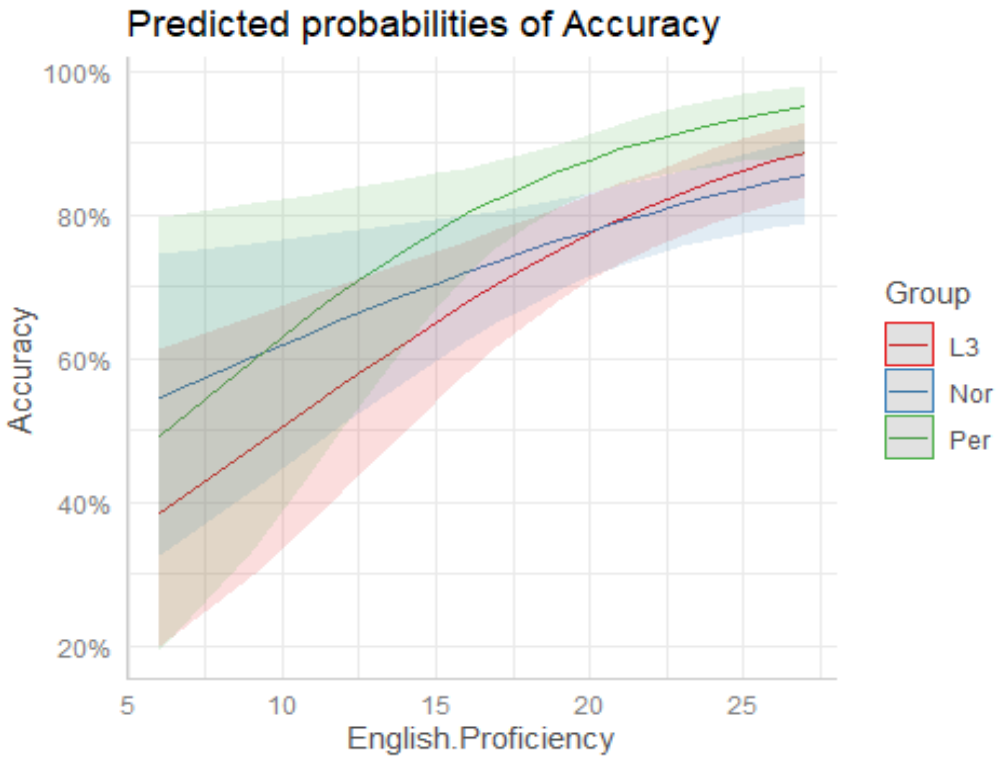


Figure 21- Predicted Probabilities of Accuracy by English Proficiency and Group

Figure 22 illustrates acceptability of the grammatical and ungrammatical sentences in each property by the three groups. To see if there were statistical differences between the acceptance of grammatical and ungrammatical sentences, two generalized linear mixed effect logistic regression models were fit, where the acceptance of the sentences was predicted by syntactic condition, group, and their interaction, with participants and sentences as random intercepts.

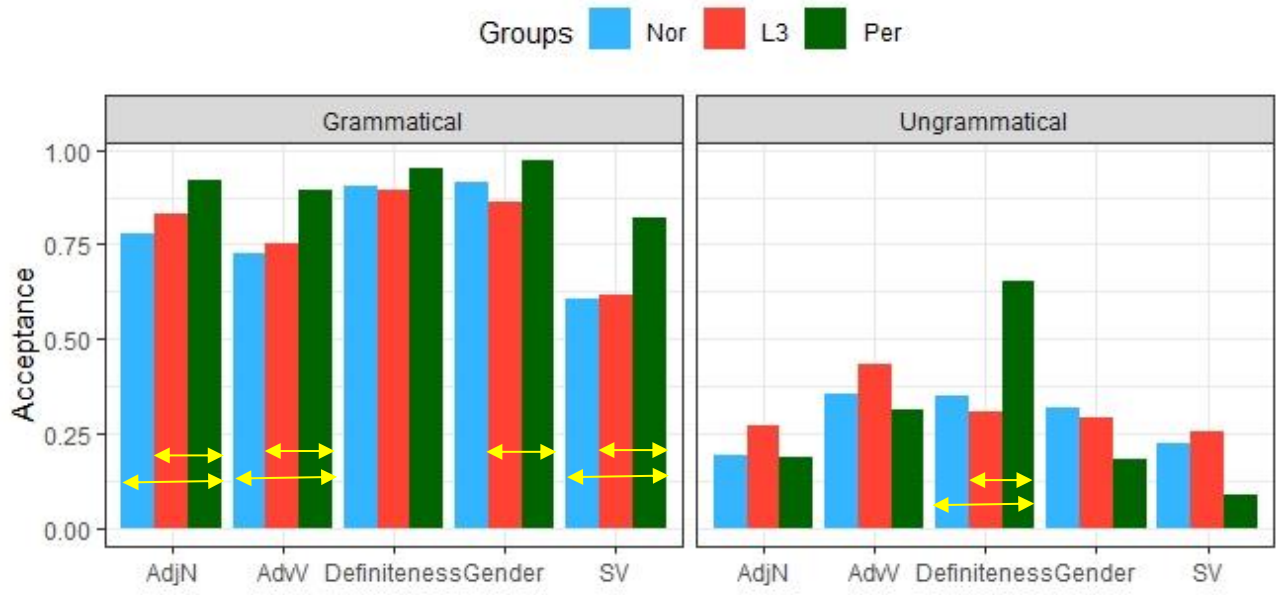


Figure 22 - Acceptability of Grammatical and Ungrammatical Sentences (significant contrasts between the groups marked with arrows)

Post-hoc pairwise comparison of the groups within properties for the grammatical condition showed that the Persian controls significantly outperformed (p-value<0.05) the L3 group in the Adj-N and gender properties and the Norwegian controls in the Adj-N property, which was not expected. There were no significant differences between the groups in the definiteness property in the grammatical condition. In the ungrammatical condition, no significant differences were found between the groups in the Adj-N and gender properties, however, the Persian controls were significantly outperformed by the other two groups in the definiteness property (p-value < 0.0001), while the L3 group and the Norwegian controls were statistically the same, which was expected.

Post-hoc pairwise comparison of the groups within properties for the grammatical condition revealed that the Persian control group significantly outperformed (p-value<0.05) both the L3 group and the Norwegian controls in the SV and Adv-V, which was expected. In the ungrammatical condition, although the groups were numerically different, the difference did not reach significance in the SV and Adv-V properties.

In sum, the results showed that the Persian control group performed more accurately than the other two groups in almost all properties, except for the definiteness property in which the L3 group had the highest accuracy. All groups were statistically the same in three out of

five properties. Significant differences between the groups were only found in the definiteness and SV properties; Persian controls outperformed the other two groups in the SV property, while the L3 and the Norwegian control groups outperformed the Persian controls in the definiteness property. The accuracy rates of the Norwegian control group and the L3 group were statistically the same in all properties. The acceptability rates of the grammatical and ungrammatical sentences confirmed further that the Persian controls were more accurate in judging the sentences as grammatical and ungrammatical in almost all properties, while the accuracy rates of the Norwegian controls and the L3 group were statistically similar. However, Persian controls were the least accurate in rejecting the ungrammatical sentences in the definiteness property, while the L3 group had the lowest rate of acceptability, which means they were the most accurate in rejecting the ungrammatical sentences in the definiteness property.

4.4 Chapter Summary

The main objective of this chapter was to see if there were significant differences between the RTs of grammatical and ungrammatical sentences in each property per group in the SPR task, and if there were significant differences between the groups. Furthermore, the results from the AJT were used in order to see if the accuracy scores of each group differed significantly from each other.

The accuracy scores of the min-AJTs in Norwegian and Persian showed that the participants in the L3 group successfully acquired four out of five properties in the background languages, however, they had not fully acquired the definiteness property.

In general, the RTs from the SPR task showed that the participants in the L3 group were relatively faster readers than the other two groups, while the Persian controls were the slowest and the Norwegian controls were in the middle in most properties, however, the groups were not statistically different. The SPR results showed further that all three groups did not differentiate significantly between the grammatical and ungrammatical conditions.

The results of the AJT showed that all groups were statistically the same in the Adj-N, Adv-V, and gender properties. The SV and definiteness property were the two critical properties, where significant differences were found between groups. In the SV property, the Persian controls were significantly more accurate than the other two groups, and the L3 and the

Norwegian control groups were statistically the same. In the definiteness property, however, the L3 group and the Norwegian controls outperformed the Persian control group, while there was no statistically significant difference between the L3 group and the Norwegian control group.

In the following chapter, I discuss the results of this study and explain further if they confirm the predictions made in chapter 3.

5 Discussion

As previously mentioned (see section 3.1), the present study addressed two main research questions:

RQ 1. What is the source of CLI? Is it only one of the previously learned languages or both?

RQ 2. Does CLI occur property by property or as a wholesale phenomenon?

These questions are answered in this chapter based on the results and findings presented in chapter 4 and the predictions drawn from the main two models of L3A introduced in chapter 2.

Based on the LPM, it was hypothesized that the source of (facilitative and non-facilitative) CLI will be both previously acquired languages, and that CLI would occur on a property-by-property basis. In contrast, it was hypothesized based on the TPM that Norwegian would be the sole source of (facilitative and non-facilitative) CLI, and CLI will occur as a wholesale phenomenon. More specific predictions based on the LPM were the following:

- For language properties where Norwegian and English pattern together, but Persian is different, it is expected that there would be a significant difference between the RTs of the grammatical and ungrammatical sentences in the L1-Norwegian group. This difference would not reach significance in the L1-Persian group, and the L3 group would differentiate more than the L1-Persian group and less than the L1-Norwegian group in the SPR task (the RTs of the L3 group would be in the middle of the two control groups in both conditions). In the AJT, the L3 group was expected to outperform the L1-Persian group and be outperformed by the L1-Norwegian group. The opposite trend was expected for language properties where Persian and English pattern together.

Furthermore, it was predicted based on the TPM:

- For language properties where Norwegian and English pattern together, but Persian is different, it was expected that the L3 and L1-Norwegian groups would differentiate between the grammatical and ungrammatical conditions significantly, while this difference would not reach significance in the L1-Persian group in the SPR task. In the

AJT, the L3 and the L1-Norwegian groups were expected to score the highest and outperform the L1-Persian group. The opposite trend was expected for language properties where Persian and English pattern together.

5.1 The SPR Task

In the SPR task, although the RTs between the grammatical and ungrammatical sentences were numerically different, none of the groups made statistically significant differences between the grammatical and ungrammatical conditions across all properties. However, the expected general trend was observed, where the RTs of the ungrammatical sentences were longer than the grammatical ones, and an effect of English proficiency (and in some cases region length) was found.

What is apparent from the results is that it is very difficult to get significant results from the SPR task. As shown in section 3.6, even the English native speakers in the pilot studies did not differentiate significantly between the grammatical and ungrammatical sentences, and the results did not change even when major modifications were made such as: 1) The design of the SPR task display was changed from linear non-cumulative to center non-cumulative to make the task more challenging. In the former display, the readers are able to see the length of each word and sentence, however, in the latter display, the readers only see one word at a time, and therefore do not know what to expect next. 2) A high number of grammatical filler sentences were added, so that the ungrammatical sentences stand out, and the readers can identify them more easily. The SPR results of the English native speakers in the pilot studies were in line with the findings of Kaan et al., (2019) and Fine et al., (2013). Both of the studies mentioned earlier were not successful in obtaining significant results from SPR tasks. In both cases, the English native speakers were tested in their native language (and compared to an L2 group in Kaan's study). The native speakers did not show an increased sensitivity toward the unexpected structures containing syntactic ambiguities, while this behavior was observed in the L2 group in Kaan's study. It was concluded that the English native speakers rapidly adapted to the unexpected structures, and therefore showed a decrease in processing difficulty.

Although the adult Persian L2 learners of English in the pilot study did differentiate between the grammatical and ungrammatical sentences in some of the properties, the three groups in the present study did not show significant sensitivity to ungrammatical sentences in

any of the properties. One possibility could be that the learners in the three groups had not acquired the properties investigated in the present study, and therefore they were not able to differentiate between the grammatical and ungrammatical conditions in the SPR task. However, as shown in section 4.2, the results of the AJT showed that all three groups were proficient enough (70% accuracy in almost all properties) to distinguish between the grammatical and ungrammatical sentences in those language properties.

It may be the case that the adaptation to ungrammatical sentences could be extended to the L2 and L3 learners in the SPR tasks[9]. Therefore, it could be argued that the participants of the present study did show sensitivity to the first ungrammatical sentences they had read, but then they adapted to the ungrammaticality. As a result, the differences in the RTs between the grammatical and ungrammatical sentences did not reach significance. It may also be possible that the regression models simply did not have enough power to show an effect of grammaticality in the groups, or the participants did not focus on the task as expected and did not do it meaningfully.

Finally, as shown in section 4.1, the Persian control group was the slowest in almost all properties in both grammatical and ungrammatical conditions, while the L3 and the Norwegian control groups were faster. Additionally, it was shown that as English proficiency increased, the RTs became longer in the Persian control group. Persian has a different writing system; its script is different from English and Norwegian, and unlike these Germanic languages, Persian is written from right to left. When two language systems have orthographic overlaps, the readers would respond more quickly in sentence reading (see Bultena et al., 2014, and Allen et al., 2021 for a detailed discussion). Thus, the orthographic differences between the language systems of the triad in the present study may be the reason why the Persian controls were the slowest in most cases, even in higher levels of English proficiency.

To sum up, despite all the effort put into making the SPR task work as expected, the present study was not successful in finding any effects of grammaticality in any of the groups. However, it was shown that English proficiency and (in some cases) region length are decisive

[9] Kaan et al., (2019, p.19) explain further that they might have obtained adaptation effects in the L2 group if they had used a long enough study.

factors in predicting RTs, so that participants with higher English proficiency had faster RTs than the participants with lower proficiency, and as the region length increased, RTs became longer, which was expected. Since no effect of grammaticality was found in the present study, the SPR predictions drawn from the TPM and LPM cannot be confirmed.

5.2 The AJT

The three groups in the present study behaved quite differently from the SPR task in the AJT. Although there were no significant differences between the RTs of grammatical and ungrammatical sentences, the AJT results showed that (see section 4.2) all groups had high accuracy scores (70% or more) in the Adj-N and gender properties, where English patterns with Norwegian, but Persian is different. The groups were also highly accurate in the Adv-V property, where English patterns with Persian , but Norwegian is different. However, the groups were only numerically different, and there were no statistically significant differences between the groups in the mentioned properties. Significant differences were found only in the definiteness and SV properties. In the definiteness property, the Norwegian controls and the L3 group significantly outperformed the Persian control group, while the Norwegian control group and the L3 group were not statistically different from each other. In the SV property, however, the Persian control group significantly outperformed the Norwegian controls and the L3 group, while there were no significant differences between the Norwegian control group and the L3 group again.

Overall, the AJT results showed that the Persian controls were more accurate than the other two groups in four out of five properties (they were only numerically different in three of the properties), while the L3 group and the Norwegian controls pattered together across all properties with no statistically significant differences.

The reason why the Persian controls were more accurate than the other two groups in most properties, even when Persian and English did not pattern together, was due to age differences, since age was found to be a significant factor in predicting accuracy in the AJT. As mentioned in section 3.5, the target age group was 11-14. However, in order to make the groups comparable in terms of their English proficiency, older participants were added to the Persian control group. The older participants in the Persian control group have had more years of English instruction at school compared to the younger participants. Consequently, they have

received more explicit instructions, especially about language properties that do not pattern with English, e.g., Adj-N and gender. Therefore, the addition of the older participants made the Persian controls generally more accurate on the various conditions than the other two groups, even though their general proficiency, measured by our test, was not higher.

However, the Persian controls were outperformed by the other two groups in the definiteness property. As mentioned in section 2.4, definiteness has been shown to be a difficult property to acquire in L2 acquisition for Persian speakers even at advanced levels, since Persian does not mark definiteness. Therefore, the results of the AJT are in line with Momenzade et al., (2014) and Joolaee et al., (2015) who found that definiteness is a challenging property to acquire for Persian learners of L2 English. Furthermore, Rajabi (2022) investigated the Bottleneck hypothesis in Persian L2 learners of English. Adj-N and gender were two of the properties which were tested, and the results showed that the Persian speakers did experience difficulty in acquiring the gender property (57% accuracy), however they were quite accurate in the Adj-N property (95%). The Persian controls in the present study were highly accurate in both Adj-N and gender properties (86% and 89% respectively), and therefore the findings are in line with Rajabi (2022) in the Adj-N property but contradict in the gender property. However, the Persian controls in the present study were older, and therefore it is not unexpected that they performed better than the L2 learners in Rajabi's study.

Moreover, it was shown in Figure 13 that as age increased, the accuracy decreased in the Norwegian control group, which was due to surprisingly low accuracy scores of a few of the 13 and 14-year-old participants in that group (see Figure 14). It seems from the results that older age groups (13 and 14-year-olds) in the Norwegian control group did not do the AJT meaningfully, since the younger age groups generally performed better than them, although the older ones had higher proficiency scores generally. Furthermore, the Norwegian controls scored exactly on the set acquisition threshold (70%) in the Adv-V property and 69% in the SV property, which indicates that these two properties were problematic for them. This resonates with findings in Westergaard (2003), who showed that younger Norwegian learners of L2 English struggled considerably with non-V2 word order.

The findings of the AJT in the two critical properties (definiteness and SV) seem to suggest that both facilitative and non-facilitative CLI come from the typologically closer

language (in this case Norwegian). It seems that the L3 group had facilitative influence from Norwegian in the definiteness property, and they had non-facilitation from Norwegian in the SV property that resulted in the L3 group patterning with the Norwegian group. Such results could be argued to support the idea that CLI occurs as a wholesale phenomenon. The Persian language remained neutral and did not affect L3 English significantly in the two critical properties. However, the TPM is a model exclusive to the initial stages and does not make any predictions for later stages of acquisition, except as a direct development of wholesale CLI. The findings could also be interpreted as going against the predictions of the LPM, since the L3 group did not have any non-facilitative or facilitative influence from Persian in the definiteness and SV properties respectively. However, the lack of such effects seems to be due to other more general factors overriding the effects of structural similarities and differences, such as proficiency and age. Furthermore, for some of the properties there seem to be ceiling effects, i.e., the property was already acquired by all groups (e.g., Adj-N word order).

As mentioned in section 2.2, most of the heritage speakers in the L3 group were simultaneous bilinguals who had been residing in Norway their whole life or had immigrated to Norway at an early age. Therefore, they all started school in Norway, where the majority language is Norwegian, and foreign language teaching occurs through the majority language Norwegian. Thus, similar to Hopp et al., (2019) and Hopp (2019), it can be argued that the participants in the L3 group were more dominant in the majority language Norwegian. Furthermore, when teaching English occurs via Norwegian, the learners would note the similarities between the two languages to a greater extent.

Thus, the findings of the AJT in the two critical properties do not necessarily go against the LPM. As discussed earlier in this section, the groups are not really comparable due to age differences and the unfortunate accuracy scores of the older participants in the Norwegian group. These factors affected the results significantly, and that makes it difficult to make any conclusions. However, it can be argued that the findings are best explained by the role of the dominant language, and therefore the research questions addressed in the present study are answered as follows: RQ1) the findings of the present study point to CLI from only one of the previously acquired languages. The CLI source for both facilitative and non-facilitative influence was Norwegian, which is probably the dominant language of the L3 learners, and the reason why L3 English seemed to be influenced only by the Norwegian language. Furthermore,

RQ2) since the Persian language was neutral and did not affect L3 English significantly, the findings do not go against CLI as a wholesale phenomenon. However, it is still a question how likely it is for wholesale CLI to occur when L3 learners are dominant in both background languages, and other factors are controlled for. Furthermore, age and English proficiency were found to be two decisive factors in predicting accuracy; generally, older participants and participants with higher English proficiency were more accurate.

5.3 Chapter Summary

In this chapter, the findings of the present study were discussed based on the previous studies and predictions drawn from L3A models.

The SPR task did not find any effect of grammaticality in the three groups, and therefore no significant differences were found between the RTs of grammatical and ungrammatical sentences in any of the properties. However, it was found that English proficiency had a significant effect on the RTs; as English proficiency increased, the RTs became faster in the L3 and Norwegian group. Due to the orthographic differences between Persian and the other two Germanic languages, the Persian controls were the slowest in reading the sentences, even at higher levels of English proficiency. Even though major modifications were made to make the task more challenging and make it easier at the same time for the participants to identify the ungrammatical sentences and distinguish them from the grammatical ones, the participants did not differentiate significantly between the grammatical and ungrammatical conditions. Therefore, it can be argued that the participants in the present study adapted to the ungrammaticality, and therefore they did not show any increased sensitivity to the ungrammatical sentences later on.

The findings of the AJT showed that the source of both facilitative and non-facilitative CLI was the Norwegian language, while the Persian language did not affect L3 English significantly. Based on the AJT findings, it can be argued that the results are best explained based on the role of the dominant language. Since the L3 learners started school in Norway, where the majority language is Norwegian, and also English is taught through Norwegian, it can be argued that the L3 learners were more dominant in Norwegian. Therefore, the L3 learners were only influenced by Norwegian since it was their dominant language. And

regarding the nature of CLI; since Persian did not affect L3 English significantly, it can be concluded that the findings do not go against CLI as a wholesale phenomenon.

6 Conclusion

The present study explored CLI in child L3A of English by Persian-Norwegian heritage speakers in Norway. Two main research questions were posed regarding the source and nature of CLI; whether the CLI source is the L1, L2 or both, and whether CLI occurs on a property-by-property basis or as a wholesale phenomenon. There are numerous previous studies of L3A, in which the source and nature of CLI have been investigated. However, CLI is still the topic of many debates within the field of L3A, due to diverging results from various studies. The predictions of the current study were drawn from the two dominant models of L3A – the TPM, a holistic basis model (Rothman 2011, 2015; Rothman et al., 2010; Rothman et al., 2019), and the LPM, a property-by-property basis model (Westergaard et al., 2017; Westergaard, 2019). The predictions of these two models were tested by the means of two main tasks: a self-paced reading (SPR) task and an acceptability judgement task (AJT).

The SPR results showed that, in general, the RTs of the ungrammatical sentences were longer than the RTs of the grammatical sentences. Compared to the Norwegian control and the L3 groups, the Persian controls were slow readers, which may be due to orthographic differences between Persian and the other two Germanic languages in the triad (see Bultena et al., 2014, and Allen et al., 2021 for a detailed discussion). The analysis of the SPR results showed further that although the RTs of the grammatical and ungrammatical conditions were numerically different, none of the groups made statistically significant differences between the grammatical and ungrammatical conditions in any of the properties. Furthermore, there were no statistically significant differences between the groups. Since no effect of grammaticality or group was found, the research questions could not be answered based on the SPR results. Overall, it was argued that the SPR results were due to adaptation effects (Kaan et al., 2019; Fine et al., 2013), in a way that the participants did show sensitivity to the first ungrammatical sentences they had read; however, they adapted to the ungrammaticality and therefore showed decreased processing difficulty. Even though major modifications were made based on the results of the several pilot studies, the difference between the RTs of the grammatical and ungrammatical conditions did not reach significance in the SPR task.

The AJT results showed that, in general, the Persian control group performed better than the other two groups in almost all properties. Nevertheless, the three groups were statistically the same in Adj-N, Adv-V, and gender properties. The SV and definiteness were the two critical

conditions, where significant difference between the groups were found. The Norwegian controls and the L3 group were outperformed by the Persian controls in the SV property, while the Persian control group was outperformed by the other two groups in the definiteness property. Thus. The differences found were in the predicted direction. The L3 group and the Norwegian control group were statistically the same across all properties. The findings of the two critical properties point to facilitative CLI from Norwegian in the definiteness property and non-facilitation from Norwegian again in the SV property. Therefore, Norwegian was found to be the sole source of CLI for the conditions tested. The L3 learners in the present study were mostly simultaneous bilinguals who started school in Norway where the majority language is Norwegian, and they had learned English via Norwegian. Therefore, it was argued according to Hopp (2019) and Hopp et al., (2019) that the L3 learners were probably more dominant in the majority language Norwegian, and this may be the reason why L3 English was only influenced by Norwegian (in addition to overall similarity). Thus, regarding the CLI source, the findings are somewhat inconclusive, but may be best explained by the role of the dominant language.

Regarding the nature of CLI, it was concluded that the findings do not go against wholesale CLI, since the L3 group and the Norwegian controls patterned together across all properties (without statistically significant differences), and the Persian language remained neutral and did not affect L3 English significantly. However, making conclusions based on the results of the present study was difficult, since there were other factors that presumably played a role, such as age differences between the groups, as well as surprising results in the Norwegian control group, where the older participants scored lower than the younger participants.

Limitations of the Present Study

The present study suffers from some limitations, mainly due to the difficulty of finding enough participants in the three groups. First, the groups were not comparable due to age differences. Second, the older participants in the Norwegian control group did not perform as expected. These two factors affected the results, and they are probably why the regression models did not show any effects of the Persian language in L3 English in the acceptability judgement task. Therefore, if there were no age differences, and Norwegian controls had performed as expected, the regression models would have had more power to show significant effects of the conditions and groups. Instead, the findings mostly showed a significant effect of age and English proficiency rather than the conditions. Additionally, ceiling effects were observed in some of the properties (e.g., the Adj-N property), which indicates all groups had already acquired those properties. Furthermore, the SPR failed to find any effects of grammaticality and group. Although we tried to cross out every possible reason that could be the reason behind that by doing several pilot studies and making major modifications to the SPR task.

Thus, future research should control for confounding factors such as age. Furthermore, the proficiency of L3 learners in the background languages should be measured more precisely, not only in the properties under investigation, but also their general proficiency. This way, the effects of dominant language may be distinguished from the effects of structural or typological similarities. Finally, future research should consider the complexity of properties under investigation in order to avoid both ceiling and floor effects.

References

Aaronson, D., & Scarborough, H. S. (1976). Performance theories for sentence coding: Some quantitative evidence. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 56-70. https://doi.org/10.1037/0096-1523.2.1.56

Aarts, B. (2011). *Oxford modern English grammar*. Oxford University Press.Hermas, A. (2010). Language acquisition as computational resetting: Verb movement in L3 initial state. *International Journal of Multilingualism*, *7*(4), 343-362. https://doi.org/10.1080/14790718.2010.487941

Adger, David (2003).Core Syntax: A Minimalist Approach. Oxford University Press: Oxford, UKand New York.

Afzali, P. (2012). Marking definiteness in Farsi and English by Farsi speaking EFL learners.

Allen, D., Conklin, K., & Miwa, K. (2021). Cross-linguistic lexical effects in different-script bilingual reading are modulated by task. International Journal of Bilingualism, 25(1), 168–188. https://doi.org/10.1177/1367006920943974

Anwyl-Irvine, A.L., Massonié J., Flitton, A., Kirkham, N.Z., Evershed, J.K. (2019). Gorilla in our midst: an online behavioural experiment builder**.** *Behavior Research Methods.* https://doi.org/10.3758/s13428-019-01237-x

Arıbaş, D. Ş., & Cele, F. (2021). Acquisition of articles in L2 and L3 English: the influence of L2 proficiency on positive transfer from L2 to L3. *Journal of multilingual and multicultural development*, *42*(1), 19-36. https://doi.org/10.1080/01434632.2019.1667364

Bardel, C., & Falk, Y. (2007). The role of the second language in third language acquisition: The case of Germanic syntax. Second Language Research, 23(4), 459–484. https://doi.org/10.1177/0267658307080557

Bardel, C.,&Falk, Y. (2012). Behind the L2 Status Factor: A neurolinguistic framework for L3 research. In J. Cabrelli Amaro, S. Flynn, & J. Rothman (Eds.), Third language acquisition in adulthood (pp. 61–78). Amsterdam: John Benjamins.

Barton, G. E., Berwick, R. C., and Ristad, E. S. (1987). Computational Complexity and Natural Language. Cambridge, MA: MIT Press.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**(1), 1–48.

Bloom, E., & Unsworth, S. (Eds.). (2010). *Experimental methods in language acquisition research* (Vol. 27). John Benjamins Publishing.

Bultena, S., Dijkstra, T., & Van Hell, J. G. (2014). Cognate effects in sentence context depend on word class, L2 proficiency, and task. *Quarterly Journal of Experimental Psychology*, *67*(6), 1214–1241. https://doi.org/10.1080/17470218.2013.853090

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

Comrie, B. (1999). Grammatical gender systems: a linguist's assessment. *Journal of Psycholinguistic research*, *28*(5), 457-466. https://doi.org/10.1023/A:1023212225540

Dąbrowska, Ewa. "Naive v. expert intuitions: An empirical study of acceptability judgments" , vol. 27, no. 1, 2010, pp. 1-23. https://doi.org/10.1515/tlir.2010.001

Enger, H., & Corbett, G. (2012). Definiteness, Gender, and Hybrids: Evidence from Norwegian Dialects. *Journal of Germanic Linguistics*, *24*(4), 287-324. https://doi.org/10.1017/S1470542712000098

Faghih, E. (1997). A Contractive Analysis of the Persian and English Definite Articles. *IRAL: International Review of Applied Linguistics in Language Teaching*, *35*(2), 127.

Falk, Y., & Bardel, C. (2011). Object pronouns in German L3 syntax: Evidence for the L2 Status Factor. Second Language Research, 27(1), 59–82. https://doi.org/10.1177/0267658310386647

Fallah, N., Jabbari, A. A., & Fazilatfar, A. M. (2016). Source(s) of syntactic cross-linguistic influence (CLI): The case of L3 acquisition of English possessives by Mazandarani–Persian bilinguals. Second Language Research, 32(2), 225–245. https://doi.org/10.1177/0267658315618009

Ferreira, F., & Henderson, J. (1990). Use of verb information during syntactic parsing: Evidence from eye tracking and word by word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 16*, 555-568. https://doi.org/10.1037/0278-7393.16.4.555

Fine AB, Jaeger TF, Farmer TA, Qian T. (2013) .Rapid Expectation Adaptation during Syntactic Comprehension. PLoS ONE 8(10): e77661. https://doi.org/10.1371/journal.pone.0077661

Flynn, S., Foley, C., & Vinnitskaya, I. (2004). The cumulative-enhancement model for language acquisition: Comparing adults' and children's patterns of development in first, second and third language acquisition of relative clauses. *International Journal of Multilingualism*, *1*(1), 3-16. https://doi.org/10.1080/14790710408668175

Geranpayeh, A. (2000). The Acquisition of the English Article System by Persian Speakers. *Edinburgh Working Papers in Applied Linguistics*, *10*, 37-51.

Ghomeshi, J. (1997). Non-projecting nouns and the ezafe: construction in Persian. *Natural Language & Linguistic Theory*, *15*(4), 729-788. https://doi.org/10.1023/A:1005886709040

Ghomeshi, J. (2003). Plural marking, indefiniteness, and the noun phrase. *Studia linguistica*, *57*(2), 47-74. https://doi.org/10.1111/1467-9582.00099

Hermas, A. (2015). The categorization of the relative complementizer phrase in third-language English: A feature re-assembly account. International Journal of Bilingualism, 19(5), 587–607. https://doi.org/10.1177/1367006914527019

Hopp, H. (2019). Cross-linguistic influence in the child third language acquisition of grammar: Sentence comprehension and production among Turkish-German and German learners of English. *International Journal of Bilingualism*, *23*(2), 567-583. https://doi.org/10.1177%2F1367006917752523

Hopp, H., Vogelbacher, M., Kieseier, T., & Thoma, D. (2019). Bilingual advantages in early foreign language learning: Effects of the minority and the majority language. *Learning and Instruction*, *61*, 99-110. https://doi.org/10.1016/j.learninstruc.2019.02.001

Jabbari, A. A., Achard-Bayle, G., & Ablali, D. (2018). Acquisition of L3 French wh-question structure by Persian-English bilinguals. *Cogent Education*, *5*(1), 1524551. https://doi.org/10.1080/2331186X.2018.1524551

Jegerski, J. (2014). Self-paced reading. In J. Jegerski & B. VanPatten (Eds.), *Research methods in second language psycholinguistics* (pp. 20-49). New York: Routledge.

Jin, F. (2009). Third Language Acquisition of Norwegian Objects: Interlanguage Transfer or L1 Influence?. In *Third language acquisition and Universal Grammar* (pp. 144-161). Multilingual Matters. https://doi.org/10.21832/9781847691323-010

Joolaee, M. G., & Ghonsooly, B. (2015). Acquisition of the article "the" by Persian speakers. *Procedia-Social and Behavioral Sciences*, *192*, 438-446. https://doi.org/10.1016/j.sbspro.2015.06.060

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological review*, *87*(4), 329. https://doi.org/10.1037/0033-295X.87.4.329

Just, M. A., Carpenter, P. A., & Wooley, J. D. (1982). Paradigms and Processes in Reading Comprehension. *Journal of Experimental Psychology: General*, *111*(2), 228-238. https://doi.org/10.1037/0096-3445.111.2.228

Kaan, E., Futch, C., Fernández Fuertes, R., Mujcinovic, S., & Álvarez De La Fuente, E. (2019). Adaptation to syntactic structures in native and nonnative sentence comprehension. *Applied Psycholinguistics*, *40*(1), 3-27. https://doi:10.1017/S0142716418000437

Kargar, A. (2019). Acquisition and accurate use of English articles by Persian speakers. *International Journal of Foreign Language Teaching and Research*, *7*(27), 87-106.

Karimi, S. (1999). Specificity effect: evidence from Persian. https://doi.org/10.1515/tlir.1999.16.2.125

Keating, G. D., & Jegerski, J. (2015). Experimental Designs in Sentence Processing Research: A Methodological Review and User's Guide. *Studies in Second Language Acquisition*, *37*(1), 1–32. http://doi.org/10.1017/S0272263114000187

Kellerman, E., & Smith, M. S. (1986). *Crosslinguistic Influence in Second Language Acquisition*. Pergamon Press.

Kupisch, T., Snape, N., & Stangen, I. (2013). Foreign language acquisition in heritage speakers. *Linguistic Superdiversity in Urban Areas: Research approaches*, *2*, 99.

Leeser , M. , Brandl , A. , & Weissglass , C . (2011). Task effects in second language sentence processing research. In P. Trofi movich & K. McDonough (Eds.), *Applying priming methods to L2 learning, teaching, and research: Insights from psycholinguistics* (pp. 179 – 198). Amsterdam, The Netherlands: Benjamins.

Leivada, E., & Westergaard, M. (2020). Acceptable ungrammatical sentences, unacceptable grammatical sentences, and the role of the cognitive parser. *Frontiers in Psychology*, *11*, 364. https://doi.org/10.3389/fpsyg.2020.00364

Lenth, R., Singman, H., Love, J., Buerkner, P., & Herve, M. (2020). R package emmeans: Estimated Marginal means, aka Least-Squares Means. R repository: https://cran.r-project.org/web/packages/emmeans/index.html

Levshina, Natalia. 2015. How to do Linguistics with R: Data exploration and statistical analysis.Amsterdam/Philadelphia: John Benjamins Publishing Company.

Lloyd-Smith, A., Gyllstad, H., & Kupisch, T. (2017). Transfer into L3 English: Global accent in German-dominant heritage speakers of Turkish. Linguistic Approaches to Bilingualism, 7(2), 131–163. https://doi.org/10.1075/lab.15013.llo

Lohndal, T., Rothman, J., Kupisch, T., & Westergaard, M. (2019). Heritage language acquisition: What it reveals and why it is important for formal linguistic theories. *Language and Linguistics Compass*, *13*(12), e12357. https://doi.org/10.1111/lnc3.12357

Marsden, E., Thompson, S., & Plonsky, L. (2018). A methodological synthesis of self-paced reading in second language research. *Applied Psycholinguistics*, *39*(5), 861-904. http://doi.org/10.1017/S0142716418000036

Mitchell, D. C., & Green, D. W. (1978). The Effects of Context and Content on Immediate Processing in Reading. Quarterly Journal of Experimental Psychology, 30(4), 609–636. https://doi.org/10.1080/14640747808400689

Momenzade, M., & Youhanaee, M. (2014). 'Number'and Article Choice: The case of Persian Learners of English. *Procedia-Social and Behavioral Sciences*, *98*, 1186-1193. https://doi.org/10.1016/j.sbspro.2014.03.533

Montrul, S. (2008). *Incomplete acquisition in bilingualism: Re-examining the age factor*. Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.39

Na Ranong, S., & Leung, Y. I. (2009). Null objects in L1 Thai–L2 English–L3 Chinese: An empiricist take on a theoretical problem. In Y.-K. I. Leung (Ed.), Third language acquisition and Universal Grammar (pp. 162–191). Bristol: Multilingual Matters. https://doi.org/10.21832/9781847691323-011

Paradis, M. (2009). Declarative and procedural determinants of second languages. Amsterdam: John Benjamins.

Polinsky, M. (2018). *Heritage Languages and their Speakers* (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press. https://doi-org.mime.uit.no/10.1017/9781107252349

Puig-Mayenco, E., González Alonso, J., & Rothman, J. (2018). A systematic review of transfer studies in third language acquisition. Second Language Research, 36(1), 31–64. https://doi.org/10.1177/0267658318809147

Rajabi, M. (2022). *L2 Acquisition of English by Persian L1 Speakers. Comparing Morphology, Syntax and Semantics The Bottleneck Hypothesis in L2 acquisition* (Master's thesis, UiT Norges arktiske universitet). https://hdl.handle.net/10037/25743

Rayner, K., Duffy, S.A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition* 14, 191–201. https://doi.org/10.3758/BF03197692

Rezaei, M. J., & Jabari, A. (2010). The Acquisition of Definiteness Feature by Persian L2 Learners of English.

Rodina, Y., & Westergaard, M. (2015). Grammatical gender in Norwegian: Language acquisition and language change. *Journal of Germanic Linguistics*, *27*(2), 145-187. https://doi.org/10.1017/S1470542714000245

Rodina, Y., & Westergaard, M. (2017). Grammatical gender in bilingual Norwegian–Russian acquisition: The role of input and transparency. *Bilingualism: Language and cognition*, *20*(1), 197-214. https://doi.org/10.1017/S1366728915000668

Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. International Journal of Bilingualism, 13(2), 155–163. https://doi.org/10.1177/1367006909339814

Rothman, J., & Cabrelli Amaro, J. (2010). What variables condition syntactic transfer? A look at the L3 initial state. *Second Language Research*, *26*(2), 189–218. https://doi.org/10.1177/0267658309349410

Rothman, J. (2011). L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research*, *27*(1), 107–127. https://doi.org/10.1177/0267658310386439

Rothman, J. (2013). Cognitive economy, non-redundancy and typological primacy in L3 acquisition: Initial stages of L3 Romance and beyond. In S. Baauw, F. Drijkoningen, L. Meroni, & M. Pinto (Eds.), *Romance Languages and Linguistic Theory* (Vol. 5, pp. 217–248). John Benjamins Publishing Company. https://doi.org/10.1075/rllt.5.11rot

Rothman, J. (2015). Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: Timing of acquisition and proficiency considered. *Bilingualism: Language and Cognition, 18*(2), 179-190. https://doi.org/10.1017/S136672891300059X

Rothman, J., Alonso, J. G., & Puig-Mayenco, E. (2019). *Third language acquisition and linguistic transfer* (Vol. 163). Cambridge University Press. https://doi.org/10.1017/9781316014660

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA http://www.rstudio.com/

Schwartz, B. D., & 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%2F026765839601200103

Sharwood Smith, M. (1983). Cross-linguistic Aspects of Second Language Acquisition. *Applied Linguistics*, *4*(3), 192–199. https://doi.org/10.1093/applin/4.3.192

Sharwood Smith, M. A. (1989). Crosslinguistic influence in language loss. In K. Hyltenstam & L. K. Obler (Eds.), *Bilingualism across the Lifespan* (1st ed., pp. 185–201). Cambridge University Press. https://doi.org/10.1017/CBO9780511611780.012

Sharwood Smith, M., & Truscott, J. (2014). *The multilingual mind: A modular processing perspective*. Cambridge University Press. https://doi.org/10.1017/CBO9781139644044

Sharwood Smith M (2019) Representing representation: Three perspectives and the case of grammatical gender. In: Slabakova R, Corbet J, Dominguez L, Dudley A, and Wellington A (eds) *Explorations in second language processing and representation*. Cambridge: Cambridge Scholars Publishing, pp. 2–40.

Sharwood Smith, M. (2021). Language transfer: a useful or pernicious concept? Second Language Research, 37(3), 409–414. https://doi.org/10.1177/0267658320941035

Slabakova, R. (2017). The scalpel model of third language acquisition. *International Journal of Bilingualism*, *21*(6), 651–665. https://doi.org/10.1177/1367006916655413

Sprouse, J., & Almeida, D. (2011). Power in acceptability judgment experiments and the reliability of data in syntax. *Ms., University of California, Irvine & Michigan State University*.

Sprouse, J., & Almeida, D. (2012). Assessing the reliability of textbook data in syntax: Adger's Core Syntax1. *Journal of Linguistics*, *48*(3), 609-652. http://doi.org/10.1017/S0022226712000011

Westergaard, M. (2003). Unlearning V2: Transfer, markedness, and the importance of input cues in the acquisition of word order in English by Norwegian children. *EUROSLA yearbook*, *3*(1), 77-101. https://doi.org/10.1075/eurosla.3.07wes

Westergaard, M. (2007). English as a mixed V2 grammar: Synchronic word order inconsistencies from the perspective of first language acquisition. *Poznań Studies in Contemporary Linguistics*, *43*(2), 107-131. https://doi.org/10.2478/v10010-007-0015-2

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

Westergaard, M. (2019). Microvariation in multilingual situations: The importance of property-by-property acquisition. Second Language Research, 37(3), 379–407. https://doi.org/10.1177/0267658319884116

Westergaard, M. (2021a). The plausibility of wholesale vs. property-by-property transfer in L3 acquisition. https://doi.org/10.1075/lab.20091.wes

Westergaard, M. (2021b). L3 acquisition and crosslinguistic influence as co-activation: Response to commentaries on the keynote 'Microvariation in multilingual situations: The importance of property-by-property acquisition'. *Second Language Research*, 02676583211007897. https://doi.org/10.1177/02676583211007897

Westergaard, M., Mitrofanova, N., Rodina, Y., & Slabakova, R. (2022). Full Transfer Potential in L3/Ln acquisition: Crosslinguistic influence as a property-by-property process. *The Cambridge Handbook of Third Language Acquisition and Processing.* Available online: https://site. uit. no/maritwestergaard/files/2022/01/WestergaardEt-al-L3Handbook2021-RevisionDec23. pdf (accessed on 16 July 2022).

Word Frequency. 2014. *Word Frequency Data: Corpus of Contemporary American English.* [ONLINE] Available at: http://www.wordfrequency.info/free.asp?s=y. [Accessed 1 October 2021]

Appendices

Appendix 1 – Information Letter and Consent Form

Dear participants,

Please read the information provided here about the experiment.

This experiment is part of a master's thesis for the 'English Acquisition and Multilingualism' master's program at UiT The Arctic University of Norway. This project explores the effects of previously learned grammars on the acquisition of English, mainly comprehension, as a third language. The main objectives of this study are as follows: First, to determine the source of this influence, whether it is L1, L2, or both. Second, to investigate variables that determine the source of influence, namely overall typological similarity and structural proximity

Participation in the project is voluntary. The study consists of a self-paced reading and acceptability judgement task in English in order to investigate the mentioned research questions. The task will be run on the web-based platform titled Gorilla Experiment Builder (www.gorilla.sc) which makes it possible to recruit participants from different parts of the world. In addition to the mentioned tasks, you should complete an English proficiency task and two other tests in Norwegian and Persian to make sure you have acquired the properties under investigation in the background languages. The researcher collects other information such as participants' age and gender, the language use and age of onset in English through a questionnaire. The whole experiment takes about 25 minutes to complete. If you choose to participate, you can withdraw your consent at any time without giving a reason. All information about you will be removed. There will be no negative consequences for you if you chose not to participate or later decide to withdraw.

We will only use your personal data for the purpose(s) specified in this information letter. We will process your personal data confidentially and in accordance with data protection legislation (the General Data Protection Regulation and Personal Data Act).

All answers and data are anonymous. The project is scheduled to end in April 2022. All the personal information will be excluded from the data after the completion of the project.

Based on an agreement with UiT The Arctic University of Norway, NSD – The Norwegian Centre for Research Data AS has assessed that the processing of personal data in this project is in accordance with data protection legislation.

Your rights

So long as you can be identified in the collected data, you have the right to:

- access the personal data that is being processed about you
- request that your personal data is deleted
- request that incorrect personal data about you is corrected/rectified
- receive a copy of your personal data (data portability), and

send a complaint to the Data Protection Officer or The Norwegian Data Protection Authority regarding the processing of your personal data

Where can I find out more?

If you have questions about the project, or want to exercise your rights, contact:

- The Arctic University of Norway (UiT) via Marit Westergaard (marit.westergaard@uit.no)
- Our Data Protection Officer: Joakim Bakkevold (personvernombud@uit.no)
- NSD – The Norwegian Centre for Research Data AS, by email: (personverntjenester@nsd.no) or by telephone: +47 53 21 15 00.

Yours sincerely,

Marit Westergaard	Farnoosh Abdollahi Dehooei
Supervisor	Student
marit.westergaard@uit.no	fde005@uit.no

Consent form

I have received and understood information about the project 'The Influence of the Previously Acquired Grammars on the Comprehension of English as a Third Language' and have been given the opportunity to ask questions. I give consent:

- to participate in the self-paced reading and acceptability judgement task
- to participate in English proficiency test and Norwegian and Persian tests
- to participate in the questionnaire
- for my child to take part in this experiment
- for information about me/myself to be published in a way that I cannot be identified
- for my personal data to be stored until the end of the project (April 2022)

☐ I have read all the information mentioned above and give consent for child to participate in this study.

(Signed by participant, date)

Appendix 2 – English Proficiency Test

Instructions: Please complete the sentences by selecting the best answer from the available answers below.

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

a. is to boil b. is boiling c. boils

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

a. there is b. is c. it is

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

a. for keeping b. to keep c. for to keep

4) In Norway people are always talking about _____.

a. a weather b. the weather c. weather

5) In Bergen _____ almost every day.

a. it rains b. there rains c. it raining

6) In the Sahara Desert there isn't _____ grass.

a. the b. some c. any

7) Some countries in Africa have _____ weather even in the cold season.

a. a warm b. the warm c. warm

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

a. coldest b. the coldest c. colder

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

a. The most b. Most of c. Most

10) Very _____ people can visit the King.

a. less b. little c. few

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

a. has won b. won c. is winning

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

a. had won b. have won c. was winning

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

a. have made him b. made him to c. made him

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

a. has b. would have c. had

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

a. both b. and c. or

16) He is very well known _____ the world.

a. all in b. all over c. in all

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

a. is believing b. are believing c. believe

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

a. from b. in c. of

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

a. had to b. must c. should

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

a. would b. will c. did

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

a. airplane b. the airplane c. an airplane

22) _____ short one.

a. quite a b. a quite c. quite

23) For many centuries men _____ to fly,

a. are trying b. try c. had tried

24) but with _____ success.

a. little b. few c. a little

25) In the 19th century a few people succeeded _____ in balloons.

a. to fly in b. flying into c. flying

26) But it wasn't until the beginning of _____ century that anybody

a. last b. next c. that

27) _____ able to fly in a machine

a. were b. is c. was

28) _____ was heavier than air,

a. who b. which c. what

29) in other words, in _____ we now call a 'plane'. The first people to achieve

a. who b. which c. what

Appendix 3 – Background Questionnaire

Please answer the questions.

How old are you?

I am ...

Please select ...

How old were you when you started learning English?

What languages do you use speaking to your mother?

What languages do you use speaking to your father?

What languages do you use speaking to your friends?

How many years have you been in Norway?

Appendix 4 – List of Test Items and Fillers

Test Item	Grammaticality	Syntactic Condition
1. My dad works in a hospital, because he is a doctor.	Grammatical	Gender
2. My sister is very little, because she is two years old.	Grammatical	Gender
3. My grandfather doesn't work, because he is very old.	Grammatical	Gender
4. My mother works in a school, because she is a teacher.	Grammatical	Gender
5. My brother can drive a car, because he is twenty years old.	Grammatical	Gender
6. Grandfather should stay home, because he is very tired.	Grammatical	Gender
7. Grandmother walks in the garden, because he likes trees a lot.	Ungrammatical	Gender
8. My mom goes to the kitchen, because he is very hungry.	Ungrammatical	Gender
9. My uncle doesn't walk home, because she has a car.	Ungrammatical	Gender
10. My aunt reads every day, because he likes books a lot.	Ungrammatical	Gender
11. Peter plays soccer, because she likes sports very much.	Ungrammatical	Gender
12. Julia helps me today, because he is free today.	Ungrammatical	Gender
13. Sara has a hat. Mary says the hat is very big.	Grammatical	Definiteness
14. There is a café here. Tim says the café is closed today.	Grammatical	Definiteness
15. My friend has a cat. Anna says the cat is very small.	Grammatical	Definiteness
16. My dad works in a bank. Mom says the bank is very big.	Grammatical	Definiteness
17. Emma has a dog. Alice says the dog is very cute.	Grammatical	Definiteness
18. There is a TV here. Dad says the TV is very old.	Grammatical	Definiteness
19. Andy has a phone. Mia says phone is very old.	Ungrammatical	Definiteness
20. My friend has a car. Sara says car is very clean.	Ungrammatical	Definiteness
21. Rosy has a dress. Anna says dress is very pretty.	Ungrammatical	Definiteness
22. Mom is writing a letter. John says letter is for grandma.	Ungrammatical	Definiteness

23. Alex has a poster. Tim says poster is very ugly.	Ungrammatical	Definiteness
24. Max has a T-shirt. Billy says T-shirt is too big.	Ungrammatical	Definiteness
25. It is lunch time. Harry wants a hot pizza for lunch today.	Grammatical	Adj-N
26. It is my book. This is a long story about two kings.	Grammatical	Adj-N
27. It is Halloween. Jim is eating sweet candies in that room.	Grammatical	Adj-N
28. It's dinner time. Anna wants a big burger to eat tonight.	Grammatical	Adj-N
29. It's sunny outside. Sara wants a cold juice to drink now.	Grammatical	Adj-N
30. We need a ruler. Mia has a long ruler on that desk.	Grammatical	Adj-N
31. I like movies. Dad plays a movie funny for me tonight.	Ungrammatical	Adj-N
32. There is a party. Mia buys a skirt short for this party.	Ungrammatical	Adj-N
33. We are at a café. Alex wants a soda cold to drink now.	Ungrammatical	Adj-N
34. It's cold today outside. John wears a jacket warm in that room.	Ungrammatical	Adj-N
35. My parents are shopping. Dad wants a shirt blue for this weekend.	Ungrammatical	Adj-N
36. Ted and John like music. Ted listens to music pop every day.	Ungrammatical	Adj-N
37. I can hear Tim and John. Today Tim laughs very loudly.	Grammatical	SV
38. Alex and Tom climbed that tree. Yesterday Alex fell very bad.	Grammatical	SV
39. My grandma was sick. Today grandma died very sadly.	Grammatical	SV
40. My friends go to the pool. Today Emma swims very well.	Grammatical	SV
41. My joke was funny. Here everyone laughed very loudly.	Grammatical	SV
42. Our team won today. Here everyone clapped very happily.	Grammatical	SV
43. Max and Tim are students. Today studies Max very carefully.	Ungrammatical	SV
44. It is Sunday. Tomorrow travels Rosy very early.	Ungrammatical	SV
45. I was very surprised . Yesterday slept Emma very late.	Ungrammatical	SV

46. Andy's project isn't finished. Today starts Andy very early.	Ungrammatical	SV
47. I am happy. Tonight arrives Alice very early.	Ungrammatical	SV
48. My family is happy. Today walks grandpa very well.	Ungrammatical	SV
49. My grandma likes sports a lot. Grandma always swims very well.	Grammatical	Adv-V
50. Sam and Sara work every day. Sam never starts very early.	Grammatical	Adv-V
51. My parents work at school. Dad often talks very slowly.	Grammatical	Adv-V
52. Tom and John are top students. Mary always studies very hard.	Grammatical	Adv-V
53. Jim and Tim are playing. Tim never screams very loudly.	Grammatical	Adv-V
54. My friends like running. Billy always runs very fast.	Grammatical	Adv-V
55. Mary and Tom go to school. Mary sleeps always very early.	Ungrammatical	Adv-V
56. My brothers love food a lot. Andy eats always very fast.	Ungrammatical	Adv-V
57. I know a lot about insects. Bees fly often very fast.	Ungrammatical	Adv-V
58. Jim and Tom are brothers. Jim cries never very loudly.	Ungrammatical	Adv-V
59. We have a meeting today. Tim arrives always very early.	Ungrammatical	Adv-V
60. Andy and Sam can sing. Andy sings often very well.	Ungrammatical	Adv-V
61. Alex is home, because school is closed today.	Grammatical	Filler
62. I want to eat out tonight. Tim is coming with me.	Grammatical	Filler
63. I want to be a nurse, because nurses can help people.	Grammatical	Filler
64. Ed and Lisa are teachers. Ed teaches math at school.	Grammatical	Filler
65. Let's watch something. This cartoon is very interesting.	Grammatical	Filler
66. I need help. This project is very difficult.	Grammatical	Filler
67. Let's visit our neighbor. Alice is new in town.	Grammatical	Filler
68. Traffic is heavy here. Traffic lights aren't working tonight.	Grammatical	Filler
69. I need to visit a dentist. My tooth hurts very much.	Grammatical	Filler

70. Ted and Ada are at the beach. Ada is taking some photos.	Grammatical	Filler
71. Summer is my favorite. Gardens are full of fruits.	Grammatical	Filler
72. It’s lunch time. Mom is making some soup.	Grammatical	Filler
73. Exams start tomorrow. Students are studying very hard.	Grammatical	Filler
74. Max and Tom are bored. Tom wants to play games.	Grammatical	Filler
75. Grandma is hungry. Grandpa is making some food.	Grammatical	Filler
76. Mia and Sam are at school. Sam is reading some books.	Grammatical	Filler
77. I study hard, because school is very important.	Grammatical	Filler
78. Kevin and Ada like art. Ada can draw very well.	Grammatical	Filler
79. Anna and Ted are outgoing. Ted visits friends every weekend.	Grammatical	Filler
80. Rosy and Billy are late, because Billy doesn’t take taxis.	Grammatical	Filler
81. Alice and Nina are sisters. Nina is older than Alice.	Grammatical	Filler
82. I have a lot of stress. My exam is this afternoon.	Grammatical	Filler
83. Mom is worried, because my brother is late.	Grammatical	Filler
84. Ted wants to visit China, because many pandas live there.	Grammatical	Filler
85. I don’t like winter, because days are very short.	Grammatical	Filler
86. My school is international. Students are from different countries.	Grammatical	Filler
87. I can’t wait for Sunday, because Sunday is Fathers' day.	Grammatical	Filler
88. I am very excited. Snow is here at last.	Grammatical	Filler
89. I like pasta. Italian food is my favorite.	Grammatical	Filler
90. We can’t make milkshake. Ted is allergic to milk.	Grammatical	Filler
91. It’s too early to sleep. Max wants to watch TV.	Grammatical	Filler
92. Alex can’t fix this phone. Tim knows more about phones.	Grammatical	Filler
93. Sara is happy, because dinner tastes very good.	Grammatical	Filler

94. I am happy, because grandma is coming next week.	Grammatical	Filler
95. Peter and Mia are home. Mia is a little sick today.	Grammatical	Filler
96. Alice buys gifts, because Christmas is in two weeks.	Grammatical	Filler
97. Mary is angry, because Ted can't cook very well.	Grammatical	Filler
98. It is summer. All students can go on vacation.	Grammatical	Filler
99. It is raining. Emma is putting on rain boots.	Grammatical	Filler
100. I feel sleepy. Coffee can help me a little.	Grammatical	Filler
101. I like drinks. Milkshakes are great in this café.	Grammatical	Filler
102. My parents like plants. Mom has many flowers at home.	Grammatical	Filler
103. I want to see Africa. Animals in Africa are very interesting.	Grammatical	Filler
104. Dad is surprised, because mom is making dinner tonight.	Grammatical	Filler
105. I have a headache. Painkillers can help me a little.	Grammatical	Filler
106. Jim and Ed are pilots. Jim can fly planes very well.	Grammatical	Filler
107. We can't go home now, because Ed wants to stay longer.	Grammatical	Filler
108. I like broccoli. Vegetables are very healthy for kids.	Grammatical	Filler
109. Let's make a cake. Anna can help us a little.	Grammatical	Filler
110. I like pets. Dogs and cats are very cute.	Grammatical	Filler
111. Australia is interesting. Many tourists go there every year.	Grammatical	Filler
112. Spring is here. All flowers bloom in this season.	Grammatical	Filler
113. Sara is sad, because Mia is still in hospital.	Grammatical	Filler
114. Rosy and Ed can play music. Ed takes music lessons every day.	Grammatical	Filler
115. I drive to work early, because traffic is very heavy here.	Grammatical	Filler
116. I go to cafes a lot. This café is famous for its cakes.	Grammatical	Filler

117. I don’t have siblings. Rosy has two sisters and one brother.	Grammatical	Filler
118. Julia is happy, because dad is coming home early today.	Grammatical	Filler
119. Billy is worried, because Tom can’t help with this project.	Grammatical	Filler
120. I am a bit tired. A cup of tea would be great.	Grammatical	Filler

Appendix 5 – List of Test Items for Norwegian Mini AJT

Test Item	Syntactic Condition	Grammaticality
1. Mora mi er lege. Hun jobber på et sykehus.	Gender	Grammatical
2. Broren min er student. Hun studerer på universitetet.	Gender	Ungrammatical
3. Jeg jobber på kontor. Kontoret er veldig stort.	Definiteness	Grammatical
4. Vennen min så et fly. Fly fløy veldig raskt på himmelen	Definiteness	Ungrammatical
1. En hvit fugl synger i treet.	Adj-N	Grammatical
2. Den sokken blå er under sengen.	Adj-N	Ungrammtical
3. I forrige uke dro faren min til Oslo.	SV	Grammatical
4. I dag vi besøker bestemor.	SV	Ungrammatical
5. Søstra mi arbeider ofte på lesesalen på universitetet.	Adv-V	Grammatical
6. Han alltid tar bussen til skolen.	Adv-V	Ungrammatical

Appendix 6 – List of Test Items for Persian Mini AJT

Test Item	Syntactic Condition	Grammaticality
۱. مداد آبی روی میز است.	Adj-N	Grammatical
۲. در خانه ی ما یک بزرگ درخت وجود دارد.	Adj-N	Ungrammtical
۳. دیروز غذا سوخت.	SV	Grammatical
۴. هفته پیش خراب شد ماشین.	SV	Ungrammatical
۵. کامپیوتر مدرسه هیچوقت کار نمی کند.	Adv-V	Grammatical
۶. مادرم کتاب میخواند همیشه.	Adv-V	Ungrammatical

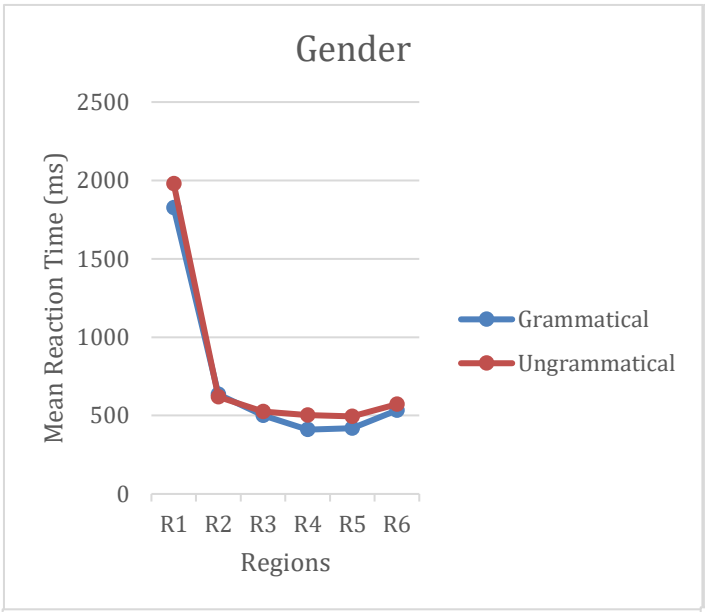
Appendix 7 – Regression Table of Proficiency Scores

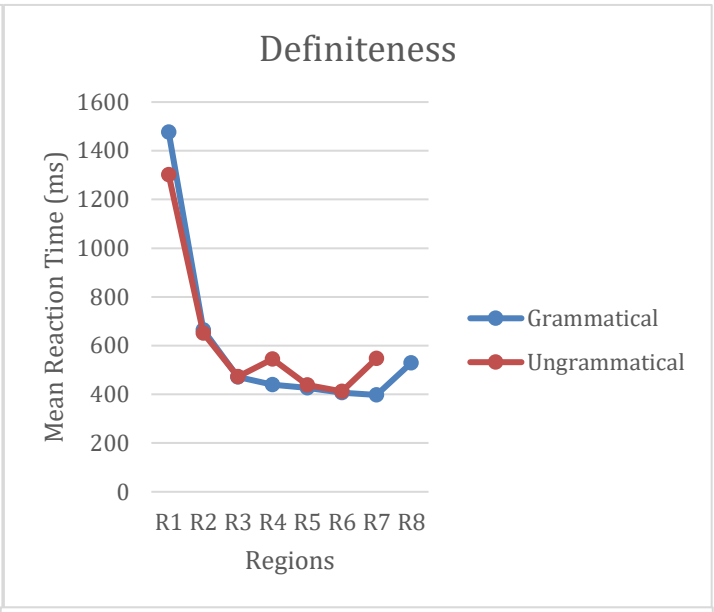
Formula: Correct ~ Group + (1 | Participant.Public.ID) + (1 | Answer)

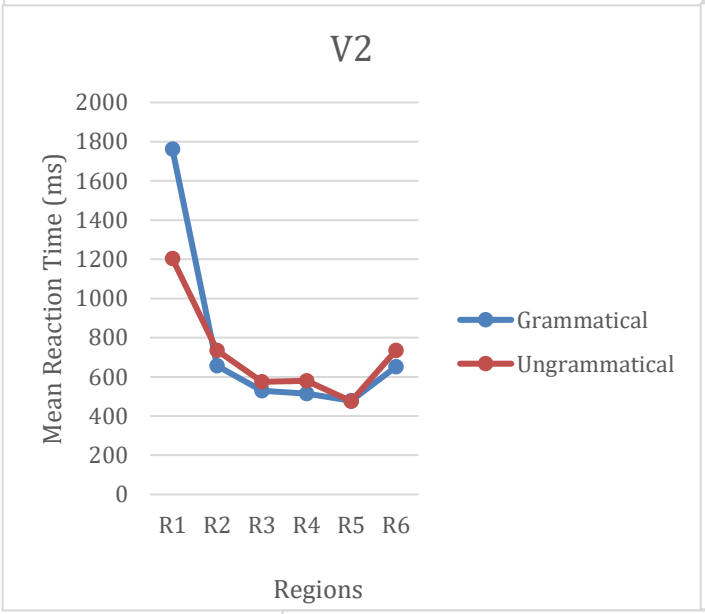
	Correct		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	3.43	1.76 – 6.72	**<0.001**
Group [Nor]	1.14	0.70 – 1.86	0.598
Group [Per]	0.67	0.41 – 1.12	0.128
Random Effects			
σ^2	3.29		
$\tau_{00\ \text{◆..Participant.Public.ID}}$	0.65		
$\tau_{00\ \text{Answer}}$	2.38		
ICC	0.48		
$N_{\text{◆..Participant.Public.ID}}$	82		
N_{Answer}	28		
Observations	2378		
Marginal R^2 / Conditional R^2	0.007 / 0.483		

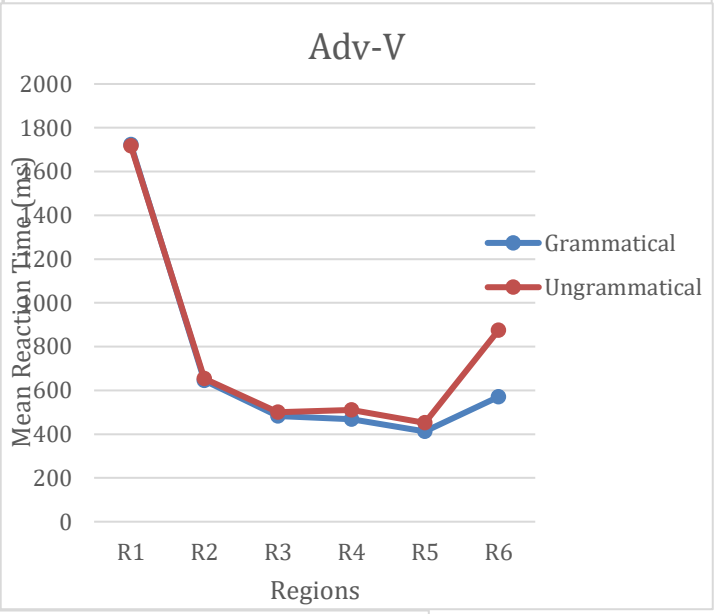
Appendix 8 – Results of the Pilot Studies

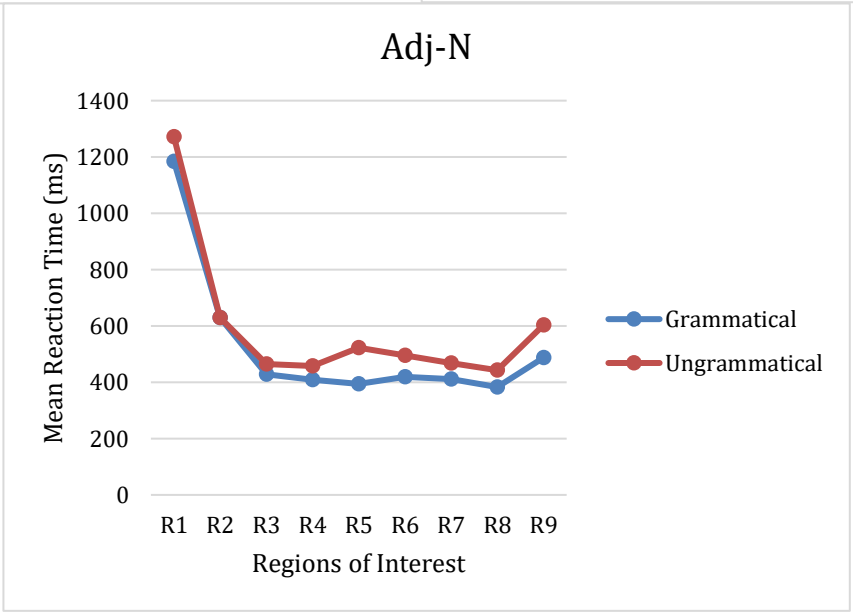
Appendix 8.1 - Results of **Pilot 1**



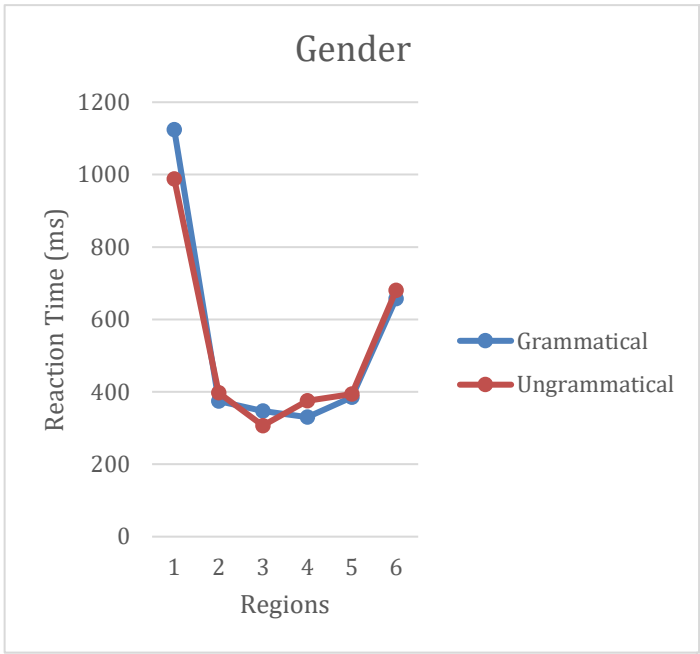


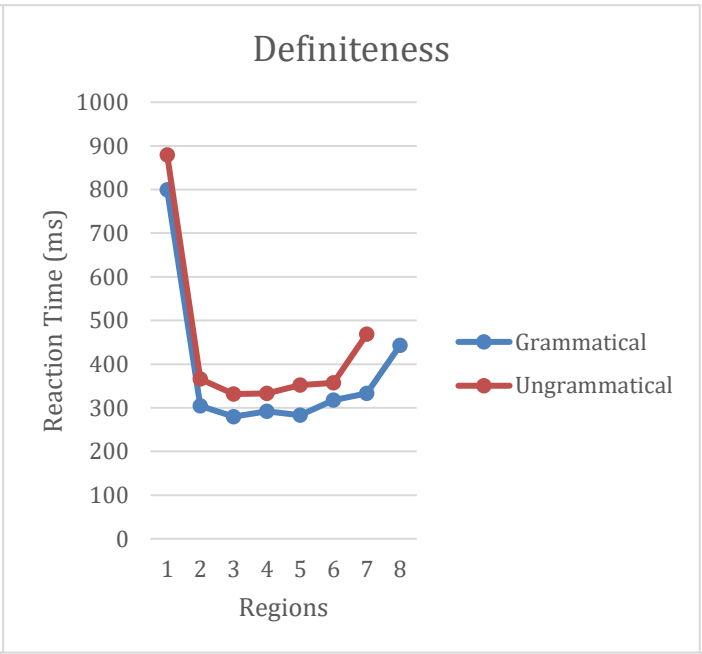


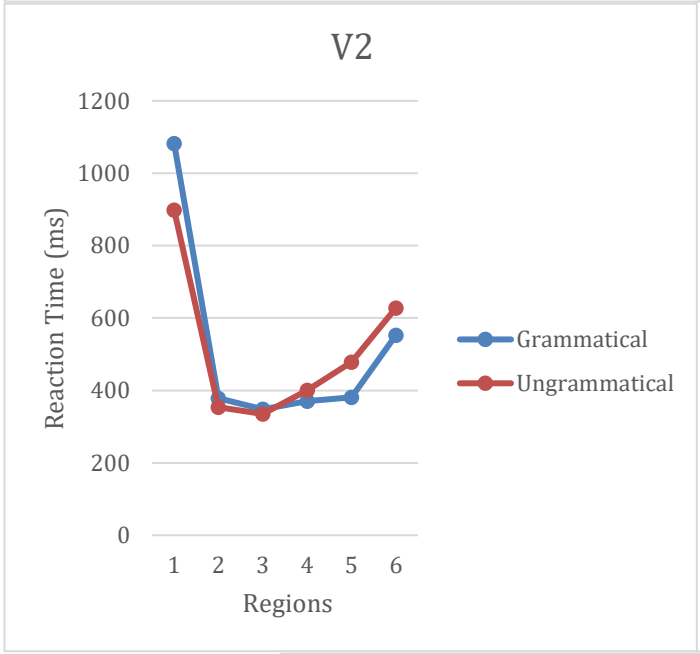


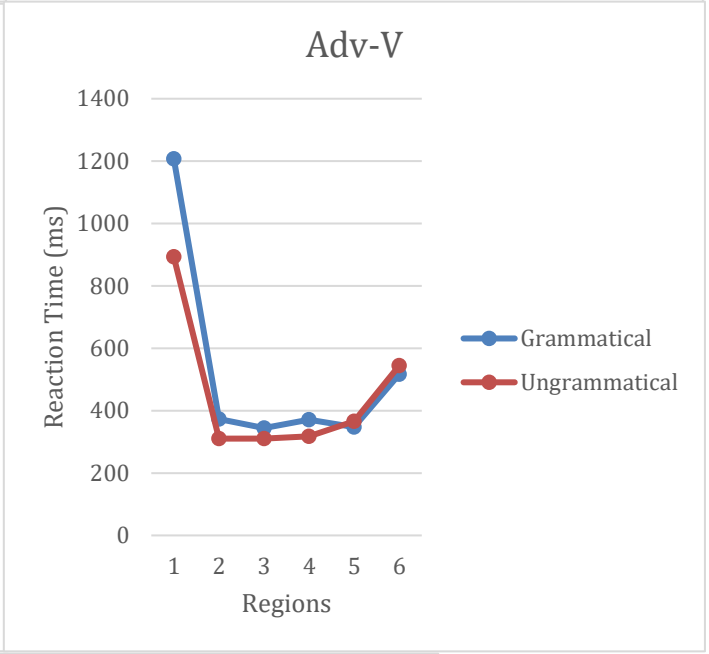


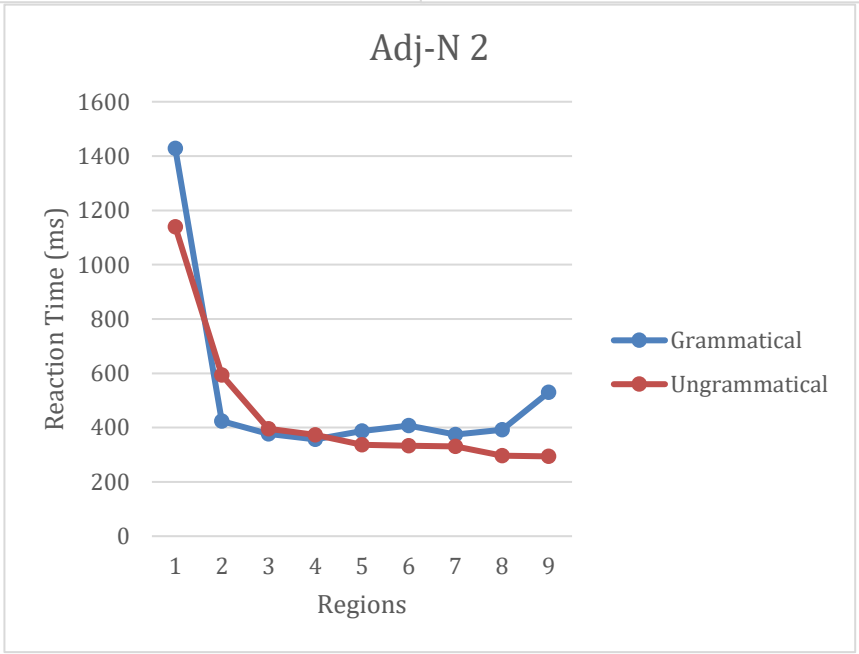
Appendix 8.2 - Results of **Pilot 2**



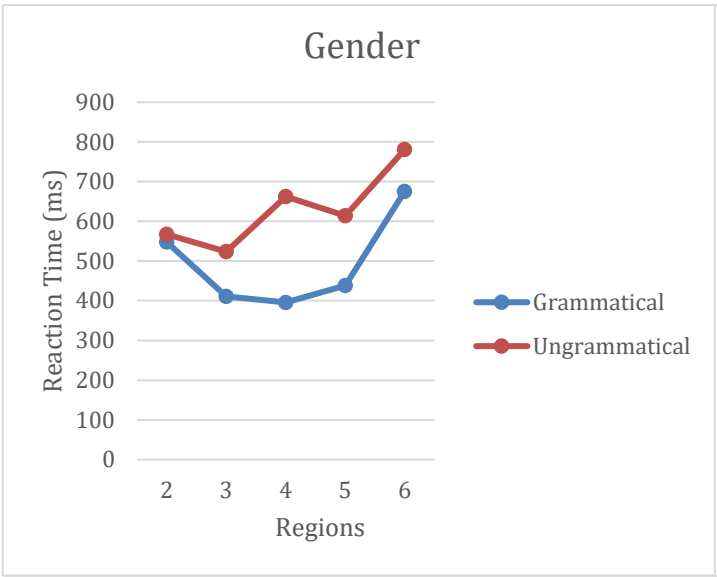


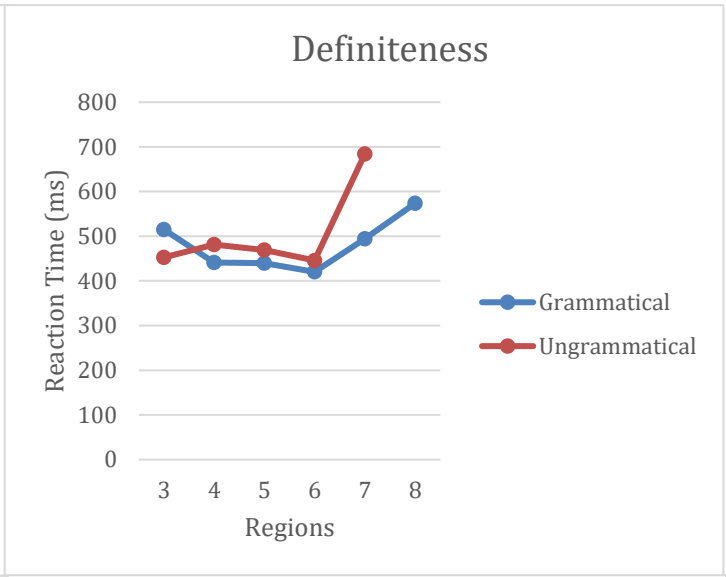


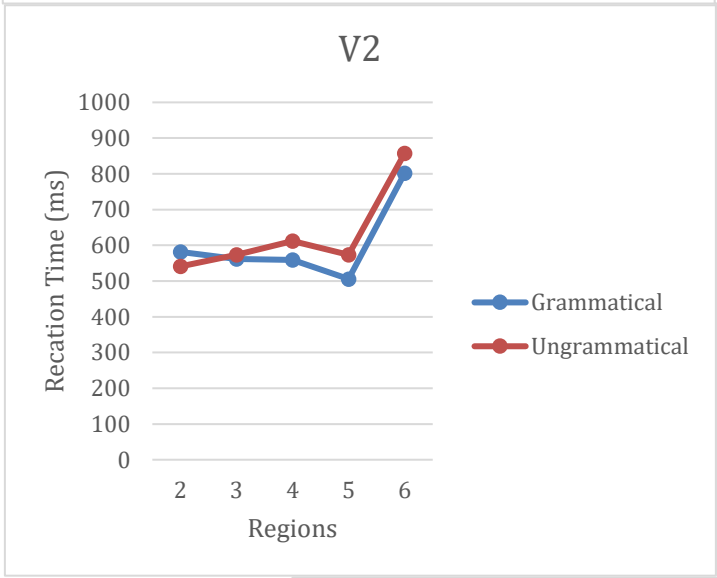


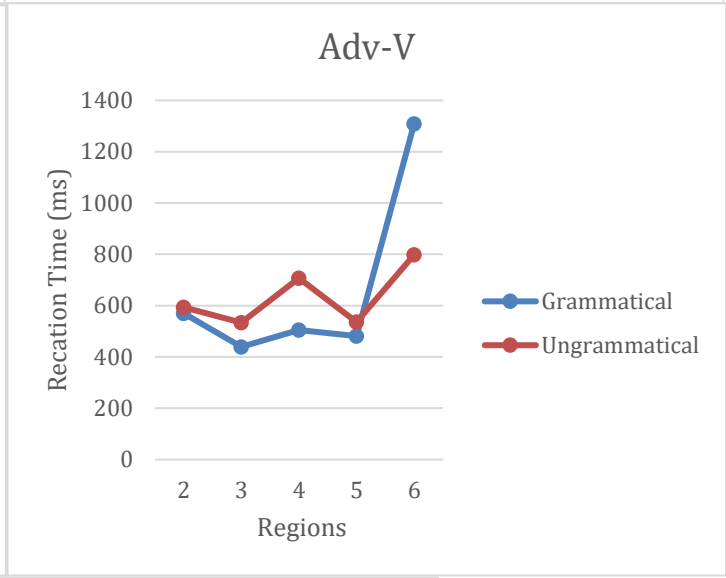


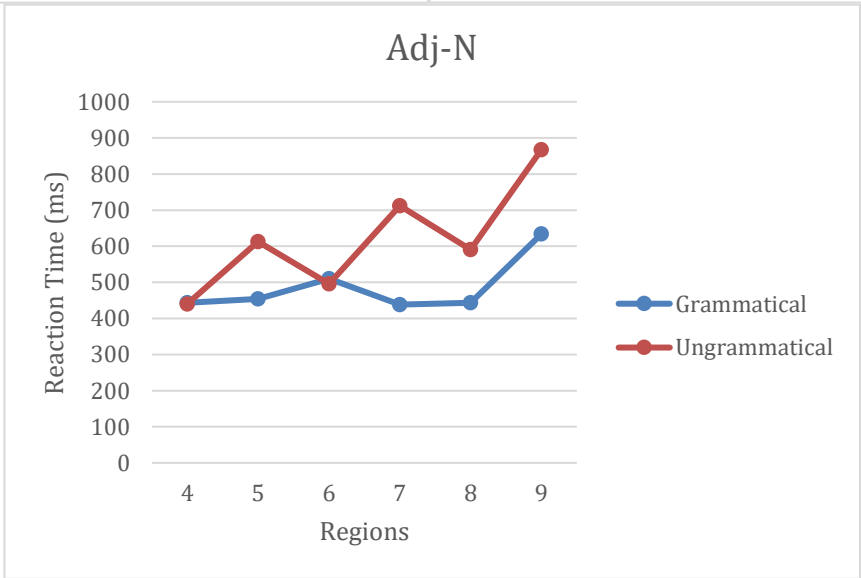
Appendix 8.3 - Results of **Pilot 3**



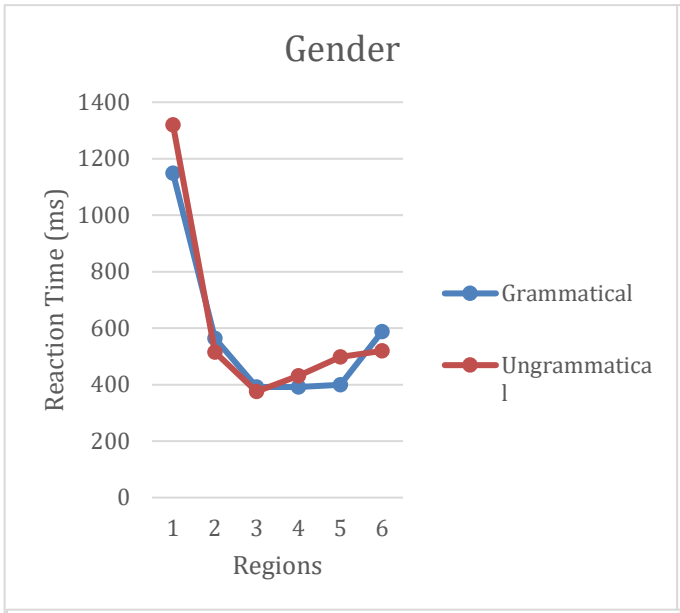


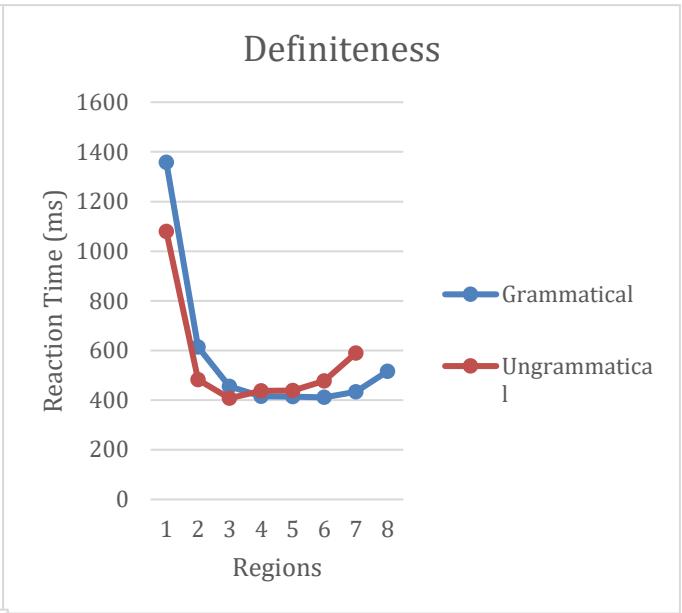


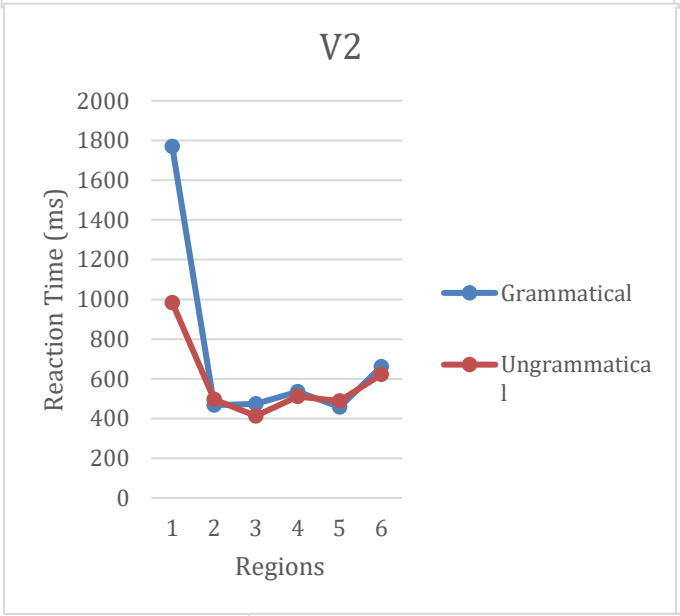


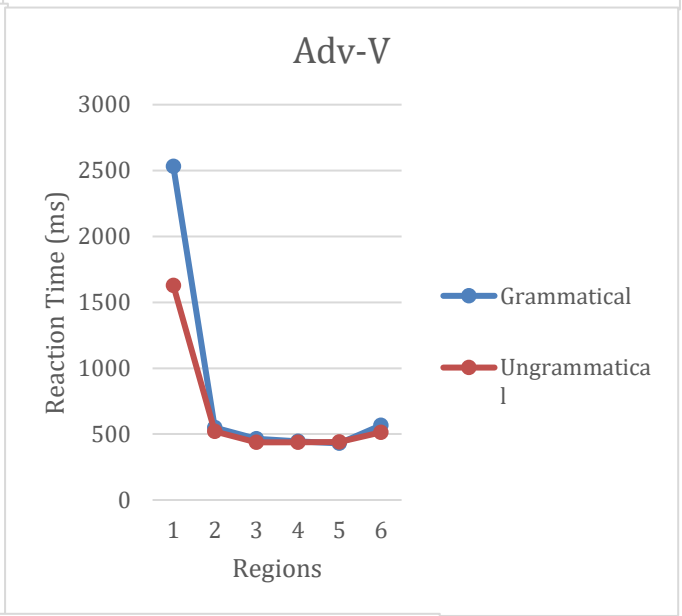


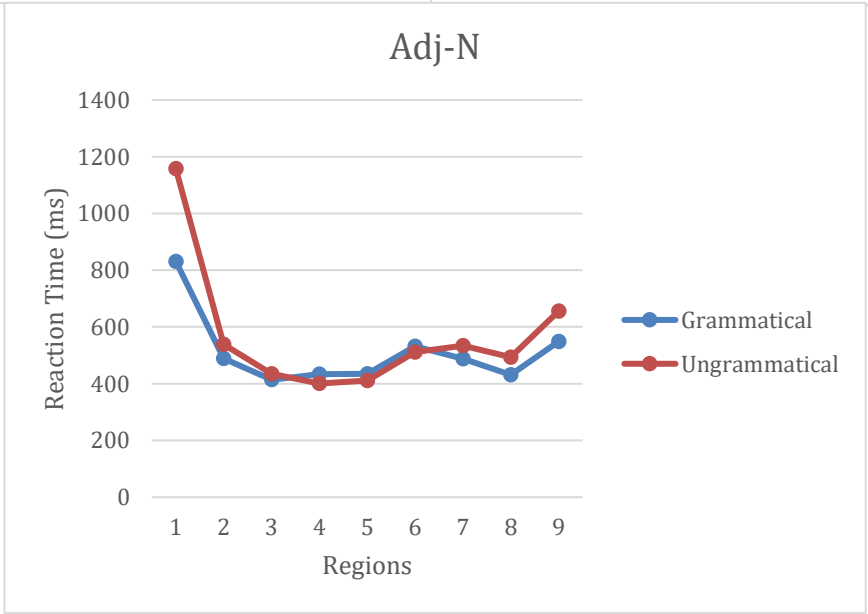
Appendix 8.4 - Results of **Pilot 4**



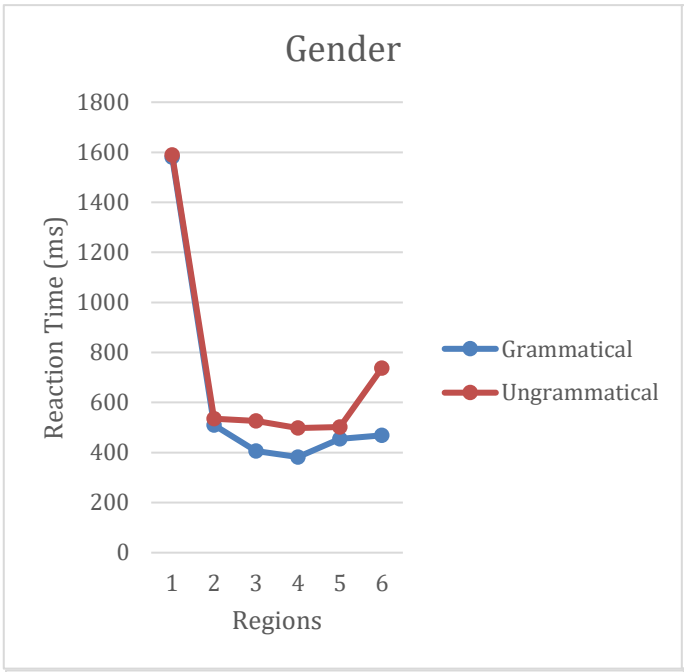


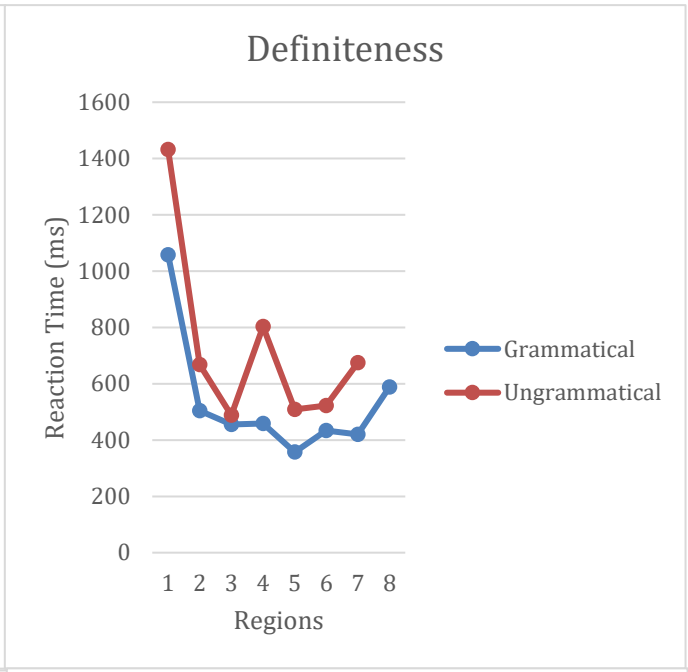


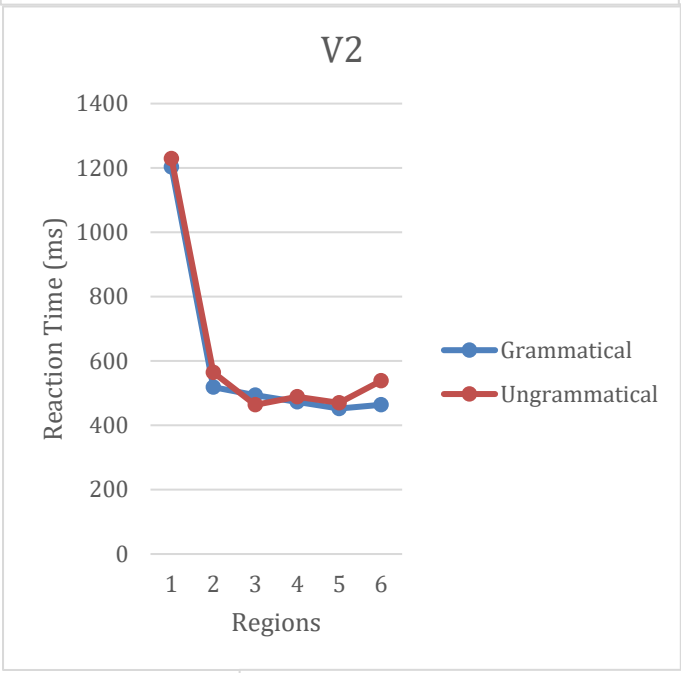


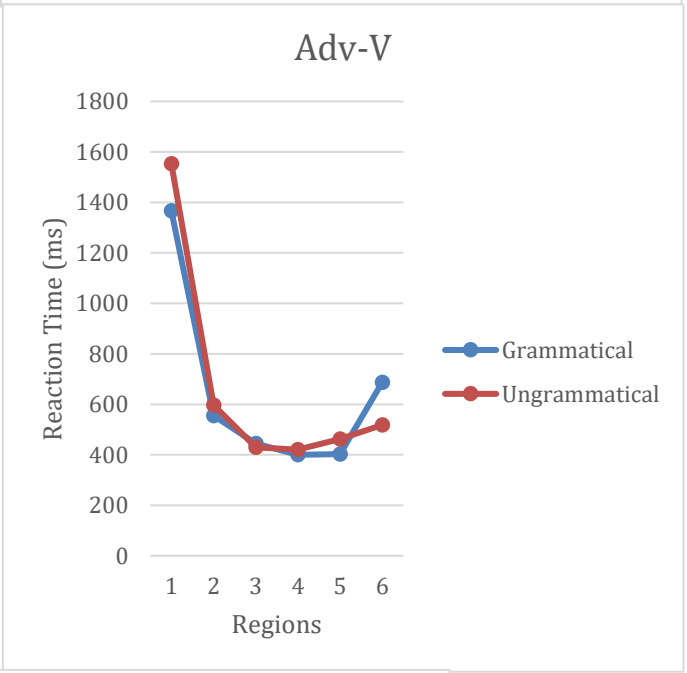


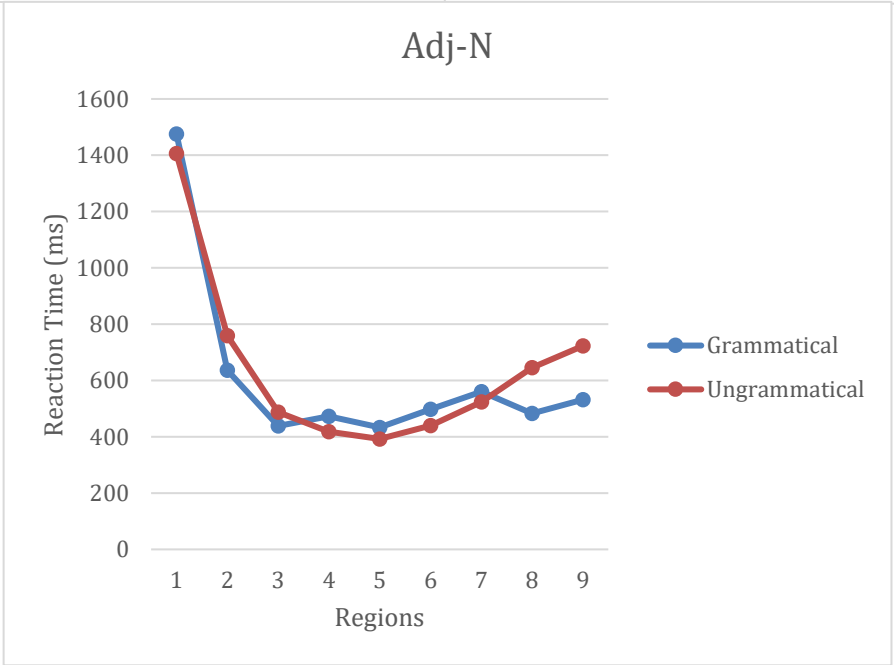
Appendix 8.5 - Results of **Pilot 5**

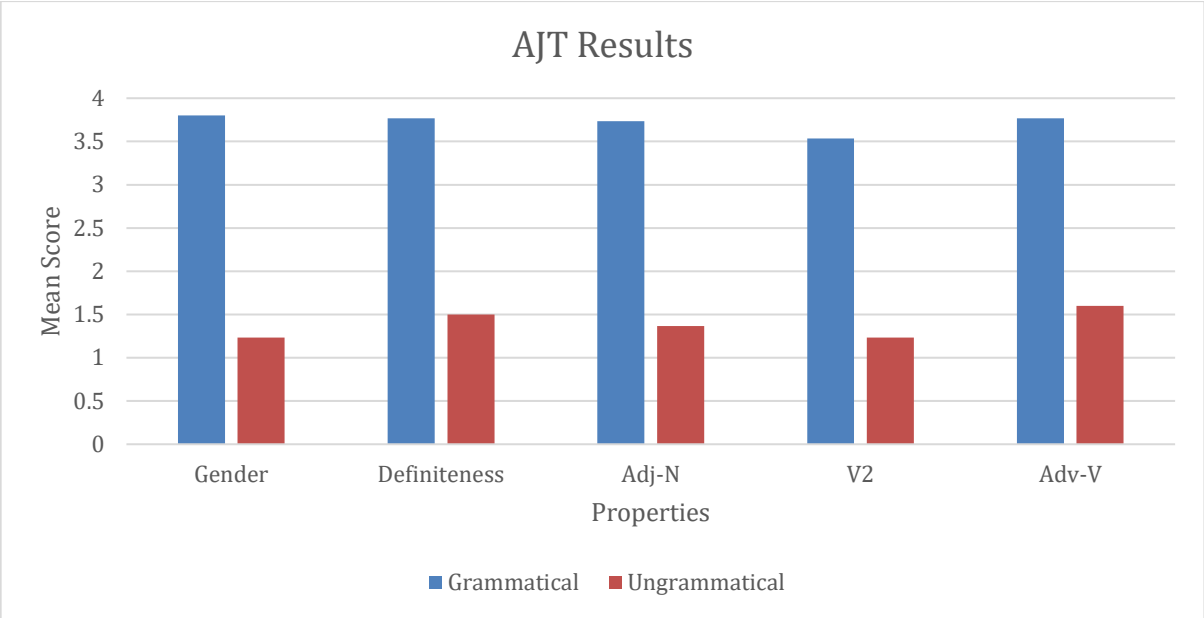











AJT Results
Mean Score
4
3.5
3
2.5
2
1.5
1
0.5
0
Gender
Definiteness
Adj-N
V2
Adv-V
Properties
Grammatical
Ungrammatical

Appendix 9 – Regression tables of the SPR task

Appendix 9.1 - Regression table of the target region in the **Adj-N** Property.

	logRT		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.74	6.21 – 7.27	**<0.001**
Group [Nor]	0.03	-0.20 – 0.27	0.772
Group [Per]	0.03	-0.22 – 0.28	0.812
Grammaticality [Ungrammatical]	0.04	-0.04 – 0.11	0.312
English Proficiency	-0.03	-0.05 – -0.00	**0.021**
region length	0.02	0.01 – 0.04	**0.010**
Group [Nor] * Grammaticality [Ungrammatical]	0.07	-0.03 – 0.17	0.194
Group [Per] * Grammaticality [Ungrammatical]	0.08	-0.02 – 0.18	0.127
Random Effects			
σ^2	0.11		
$\tau_{00\ Participant.Public.ID}$	0.19		
$\tau_{00\ Sentence}$	0.00		
ICC	0.64		
$N_{Participant.Public.ID}$	82		
$N_{Sentence}$	12		
Observations	954		
Marginal R^2 / Conditional R^2	0.059 / 0.659		

Appendix 9.2 - Regression table of the target + spillover regions in the **Adj-N** Property.

	LogSumRT		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	7.86	7.34 – 8.38	**<0.001**
Group [Nor]	0.05	-0.17 – 0.28	0.642
Group [Per]	0.06	-0.18 – 0.30	0.635
Grammaticality [Ungrammatical]	0.02	-0.05 – 0.08	0.625
English Proficiency	-0.03	-0.05 – -0.00	**0.017**
SumRegLength	0.01	-0.00 – 0.02	0.166
Group [Nor] * Grammaticality [Ungrammatical]	0.02	-0.07 – 0.11	0.636
Group [Per] * Grammaticality [Ungrammatical]	0.03	-0.06 – 0.13	0.462
Random Effects			
σ^2	0.08		
$\tau_{00\ Participant.Public.ID}$	0.17		
$\tau_{00\ Sentence}$	0.00		
ICC	0.68		
$N_{Participant.Public.ID}$	82		
$N_{Sentence}$	12		
Observations	911		
Marginal R^2 / Conditional R^2	0.055 / 0.699		

Appendix 9.3 - Regression table of English proficiency in the **Adj-N** property.

	logRT		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	7.75	7.01 – 8.48	**<0.001**
English Proficiency	-0.07	-0.11 – -0.04	**<0.001**
Group [Nor]	-1.31	-2.33 – -0.29	**0.012**
Group [Per]	-1.74	-3.06 – -0.42	**0.010**
English Proficiency * Group [Nor]	0.07	0.02 – 0.11	**0.007**
English Proficiency * Group [Per]	0.09	0.02 – 0.16	**0.008**
Random Effects			
σ^2	0.11		
$\tau_{00\ Participant.Public.ID}$	0.16		
$\tau_{00\ Sentence}$	0.00		
ICC	0.60		
$N_{Participant.Public.ID}$	82		
$N_{Sentence}$	12		
Observations	2850		
Marginal R^2 / Conditional R^2	0.114 / 0.642		

Appendix 9.4 - Regression table of the target region in the **definiteness** property.

	logRT		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.69	6.18 – 7.21	**<0.001**
Group [Nor]	0.07	-0.16 – 0.30	0.535
Group [Per]	0.04	-0.20 – 0.28	0.741
Grammaticality [Ungrammatical]	0.08	-0.01 – 0.16	0.092
English Proficiency	-0.03	-0.05 – -0.00	**0.030**
region length	0.01	-0.01 – 0.03	0.337
Group [Nor] * Grammaticality [Ungrammatical]	-0.01	-0.10 – 0.09	0.918
Group [Per] * Grammaticality [Ungrammatical]	0.02	-0.08 – 0.12	0.740
Random Effects			
σ^2	0.10		
$\tau_{00\ Participant.Public.ID}$	0.18		
$\tau_{00\ Sentence}$	0.00		
ICC	0.63		
$N_{Participant.Public.ID}$	82		
$N_{Sentence}$	12		
Observations	954		
Marginal R^2 / Conditional R^2	0.047 / 0.649		

Appendix 9.5 - Regression table of the target + spillover regions in the **definiteness** property.

	LogSumRT		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	7.68	7.18 – 8.18	**<0.001**
Group [Nor]	0.07	-0.14 – 0.29	0.499
Group [Per]	0.02	-0.20 – 0.25	0.855
Grammaticality [Ungrammatical]	0.01	-0.06 – 0.09	0.713
English Proficiency	-0.02	-0.04 – -0.00	**0.040**
SumRegLength	0.01	-0.01 – 0.03	0.243
Group [Nor] * Grammaticality [Ungrammatical]	0.02	-0.07 – 0.10	0.651
Group [Per] * Grammaticality [Ungrammatical]	0.03	-0.05 – 0.12	0.448
Random Effects			
σ^2	0.08		
$\tau_{00\ Participant.Public.ID}$	0.15		
$\tau_{00\ Sentence}$	0.00		
ICC	0.67		
$N_{Participant.Public.ID}$	82		
$N_{Sentence}$	12		
Observations	931		
Marginal R^2 / Conditional R^2	0.041 / 0.683		

Appendix 9.6 - Regression table of English proficiency in the **definiteness** property.

	logRT		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	7.75	7.01 – 8.48	**<0.001**
English Proficiency	-0.07	-0.11 – -0.04	**<0.001**
Group [Nor]	-1.31	-2.33 – -0.29	**0.012**
Group [Per]	-1.74	-3.06 – -0.42	**0.010**
English Proficiency * Group [Nor]	0.07	0.02 – 0.11	**0.007**
English Proficiency * Group [Per]	0.09	0.02 – 0.16	**0.008**
Random Effects			
σ^2	0.11		
$\tau_{00\ Participant.Public.ID}$	0.16		
$\tau_{00\ Sentence}$	0.00		
ICC	0.60		
$N_{Participant.Public.ID}$	82		
$N_{Sentence}$	12		
Observations	2850		
Marginal R^2 / Conditional R^2	0.114 / 0.642		

Appendix 9.7 - Regression table of the target region in the **SV** property.

	logRT		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.90	6.30 – 7.50	**<0.001**
Group [Nor]	0.03	-0.23 – 0.29	0.835
Group [Per]	0.04	-0.24 – 0.32	0.774
Grammaticality [Ungrammatical]	0.02	-0.07 – 0.10	0.717
English Proficiency	-0.03	-0.06 – -0.01	**0.019**
region length	0.02	-0.01 – 0.05	0.117
Group [Nor] * Grammaticality [Ungrammatical]	0.00	-0.10 – 0.11	0.955
Group [Per] * Grammaticality [Ungrammatical]	-0.00	-0.11 – 0.11	0.951
Random Effects			
σ^2	0.12		
$\tau_{00\ Participant.Public.ID}$	0.24		
$\tau_{00\ Sentence}$	0.00		
ICC	0.67		
$N_{Participant.Public.ID}$	82		
$N_{Sentence}$	12		
Observations	935		
Marginal R^2 / Conditional R^2	0.051 / 0.688		

Appendix 9.8 - Regression table of the target + spillover regions in the **SV** property.

	LogSumRT		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	7.85	7.27 – 8.42	**<0.001**
Group [Nor]	0.06	-0.19 – 0.30	0.652
Group [Per]	0.01	-0.25 – 0.28	0.914
Grammaticality [Ungrammatical]	0.06	-0.00 – 0.13	0.055
English Proficiency	-0.03	-0.06 – -0.01	**0.019**
SumRegLength	0.01	0.00 – 0.02	**0.007**
Group [Nor] * Grammaticality [Ungrammatical]	-0.04	-0.13 – 0.05	0.403
Group [Per] * Grammaticality [Ungrammatical]	0.01	-0.09 – 0.10	0.884
Random Effects			
σ^2	0.08		
$\tau_{00\ Participant.Public.ID}$	0.21		
$\tau_{00\ Sentence}$	0.00		
ICC	0.73		
$N_{Participant.Public.ID}$	82		
$N_{Sentence}$	12		
Observations	905		
Marginal R^2 / Conditional R^2	0.053 / 0.741		

Appendix 9.9 - Regression table of English proficiency in the **SV** property.

	logRT		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	7.75	7.01 – 8.48	**<0.001**
English Proficiency	-0.07	-0.11 – -0.04	**<0.001**
Group [Nor]	-1.31	-2.33 – -0.29	**0.012**
Group [Per]	-1.74	-3.06 – -0.42	**0.010**
English Proficiency * Group [Nor]	0.07	0.02 – 0.11	**0.007**
English Proficiency * Group [Per]	0.09	0.02 – 0.16	**0.008**
Random Effects			
σ^2	0.11		
$\tau_{00\ Participant.Public.ID}$	0.16		
$\tau_{00\ Sentence}$	0.00		
ICC	0.60		
$N_{Participant.Public.ID}$	82		
$N_{Sentence}$	12		
Observations	2850		
Marginal R^2 / Conditional R^2	0.114 / 0.642		

Appendix 9.10 - Regression table of the target region in the **Adv-V** property.

	logRT		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.84	6.22 – 7.46	**<0.001**
Group [Nor]	0.04	-0.19 – 0.28	0.734
Group [Per]	0.11	-0.14 – 0.36	0.404
Grammaticality [Ungrammatical]	-0.03	-0.11 – 0.06	0.530
English Proficiency	-0.03	-0.05 – -0.00	**0.019**
region length	0.01	-0.05 – 0.07	0.729
Group [Nor] * Grammaticality [Ungrammatical]	0.06	-0.04 – 0.16	0.217
Group [Per] * Grammaticality [Ungrammatical]	-0.02	-0.12 – 0.08	0.701
Random Effects			
σ^2	0.11		
$\tau_{00\ Participant.Public.ID}$	0.19		
$\tau_{00\ Sentence}$	0.00		
ICC	0.64		
$N_{Participant.Public.ID}$	82		
$N_{Sentence}$	12		
Observations	955		
Marginal R^2 / Conditional R^2	0.053 / 0.659		

Appendix 9.11 - Regression table of the target + spillover regions in the **Adv-V** property.

	LogSumRT		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	7.87	7.31 – 8.44	**<0.001**
Group [Nor]	0.04	-0.18 – 0.26	0.717
Group [Per]	0.10	-0.14 – 0.33	0.426
Grammaticality [Ungrammatical]	0.00	-0.07 – 0.07	0.999
English Proficiency	-0.03	-0.05 – -0.01	**0.010**
SumRegLength	0.01	-0.01 – 0.03	0.256
Group [Nor] * Grammaticality [Ungrammatical]	0.04	-0.04 – 0.13	0.351
Group [Per] * Grammaticality [Ungrammatical]	-0.05	-0.14 – 0.04	0.295
Random Effects			
σ^2	0.08		
$\tau_{00\ Participant.Public.ID}$	0.17		
$\tau_{00\ Sentence}$	0.00		
ICC	0.70		
$N_{Participant.Public.ID}$	82		
$N_{Sentence}$	12		
Observations	922		
Marginal R^2 / Conditional R^2	0.063 / 0.715		

Appendix 9.12 - Regression table of English proficiency in the **Adv-V** property.

	logRT		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	7.75	7.01 – 8.48	**<0.001**
English Proficiency	-0.07	-0.11 – -0.04	**<0.001**
Group [Nor]	-1.31	-2.33 – -0.29	**0.012**
Group [Per]	-1.74	-3.06 – -0.42	**0.010**
English Proficiency * Group [Nor]	0.07	0.02 – 0.11	**0.007**
English Proficiency * Group [Per]	0.09	0.02 – 0.16	**0.008**
Random Effects			
σ^2	0.11		
$\tau_{00\ Participant.Public.ID}$	0.16		
$\tau_{00\ Sentence}$	0.00		
ICC	0.60		
$N_{Participant.Public.ID}$	82		
$N_{Sentence}$	12		
Observations	2850		
Marginal R^2 / Conditional R^2	0.114 / 0.642		

Appendix 9.13 - Regression table of the target region in the **gender** property.

	logRT		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.66	6.15 – 7.17	**<0.001**
Group [Nor]	0.10	-0.12 – 0.33	0.369
Group [Per]	0.05	-0.19 – 0.28	0.709
Grammaticality [Ungrammatical]	0.00	0.01 – 0.14	0.202
English Proficiency	-0.02	-0.05 – -0.00	**0.033**
region length	0.01	-0.03 – 0.06	0.488
Group [Nor] * Grammaticality [Ungrammatical]	-0.01	-0.10 – 0.09	0.880
Group [Per] * Grammaticality [Ungrammatical]	0.01	-0.09 – 0.10	0.904
Random Effects			
σ^2	0.09		
$\tau_{00\ Participant.Public.ID}$	0.17		
$\tau_{00\ Sentence}$	0.00		
ICC	0.64		
$N_{Participant.Public.ID}$	82		
$N_{Sentence}$	12		
Observations	949		
Marginal R^2 / Conditional R^2	0.048 / 0.662		

Appendix 9.14 - Regression table of the target + spillover regions in the **gender** property.

	LogSumRT		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	7.69	7.20 – 8.17	**<0.001**
Group [Nor]	0.08	-0.13 – 0.30	0.458
Group [Per]	0.06	-0.16 – 0.29	0.580
Grammaticality [Ungrammatical]	0.00	0.02 – 0.14	0.095
English Proficiency	-0.02	-0.05 – -0.00	**0.036**
SumRegLength	0.01	0.00 – 0.02	**0.031**
Group [Nor] * Grammaticality [Ungrammatical]	0.01	-0.07 – 0.10	0.771
Group [Per] * Grammaticality [Ungrammatical]	0.01	-0.07 – 0.10	0.760
Random Effects			
σ^2	0.07		
$\tau_{00\ Participant.Public.ID}$	0.16		
$\tau_{00\ Sentence}$	0.00		
ICC	0.70		
$N_{Participant.Public.ID}$	82		
$N_{Sentence}$	12		
Observations	910		
Marginal R^2 / Conditional R^2	0.059 / 0.713		

Appendix 9.15 - Regression table of English proficiency in the **gender** property.

	logRT		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	7.75	7.01 – 8.48	**<0.001**
English Proficiency	-0.07	-0.11 – -0.04	**<0.001**
Group [Nor]	-1.31	-2.33 – -0.29	**0.012**
Group [Per]	-1.74	-3.06 – -0.42	**0.010**
English Proficiency * Group [Nor]	0.07	0.02 – 0.11	**0.007**
English Proficiency * Group [Per]	0.09	0.02 – 0.16	**0.008**
Random Effects			
σ^2	0.11		
$\tau_{00\ Participant.Public.ID}$	0.16		
$\tau_{00\ Sentence}$	0.00		
ICC	0.60		
$N_{Participant.Public.ID}$	82		
$N_{Sentence}$	12		
Observations	2850		
Marginal R^2 / Conditional R^2	0.114 / 0.642		

Appendix 10 – Regression tables of the AJT

Appendix 10.1 - Formula: Accuracy ~ 1 + Syntactic.Condition * Group + age + (1 | Participant.Public.ID) + (1 | sentences)

	Accuracy		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.88	0.23 – 3.36	0.848
Syntactic Condition [AdvV]	0.50	0.24 – 1.04	0.065
Syntactic Condition [Definiteness]	1.33	0.62 – 2.84	0.460
Syntactic Condition [Gender]	1.10	0.52 – 2.34	0.795
Syntactic Condition [SV]	0.54	0.26 – 1.14	0.105
Group [Nor]	1.11	0.64 – 1.91	0.712
Group [Per]	1.65	0.85 – 3.20	0.142
age	1.14	1.04 – 1.26	**0.008**
Syntactic Condition [AdvV] * Group [Nor]	1.11	0.66 – 1.88	0.689
Syntactic Condition [Definiteness] * Group [Nor]	0.87	0.49 – 1.52	0.616
Syntactic Condition [Gender] * Group [Nor]	1.02	0.59 – 1.79	0.936
Syntactic Condition [SV] * Group [Nor]	0.91	0.54 – 1.53	0.715
Syntactic Condition [AdvV] * Group [Per]	1.02	0.55 – 1.88	0.959
Syntactic Condition [Definiteness] * Group [Per]	0.17	0.09 – 0.32	<**0.001**
Syntactic Condition [Gender] * Group [Per]	1.21	0.62 – 2.38	0.576

Syntactic Condition [SV] * Group [Per]	1.66	0.87 – 3.16	0.126

Random Effects

σ^2	3.29
$\tau_{00\ Participant.Public.ID}$	0.52
$\tau_{00\ sentences}$	0.64
ICC	0.26
$N_{\ Participant.Public.ID}$	80
$N_{\ sentences}$	60
Observations	4723
Marginal R^2 / Conditional R^2	0.072 / 0.315

Appendix 10.2 – Post-hoc pairwise comparison of Appendix 14.1

```
## $contrasts
## Syntactic.Condition = AdjN:
##  contrast  odds.ratio    SE  df null z.ratio p.value
##  L3 / Nor       0.970 0.271 Inf    1  -0.109  0.9934
##  L3 / Per       0.588 0.203 Inf    1  -1.538  0.2731
##  Nor / Per      0.606 0.207 Inf    1  -1.467  0.3070
##
## Syntactic.Condition = AdvV:
##  contrast  odds.ratio    SE  df null z.ratio p.value
##  L3 / Nor       0.875 0.234 Inf    1  -0.498  0.8722
##  L3 / Per       0.579 0.188 Inf    1  -1.688  0.2097
##  Nor / Per      0.661 0.212 Inf    1  -1.294  0.3985
##
## Syntactic.Condition = Definiteness:
##  contrast  odds.ratio    SE  df null z.ratio p.value
##  L3 / Nor       1.173 0.337 Inf    1   0.556  0.8433
##  L3 / Per       3.444 1.130 Inf    1   3.770  0.0005
##  Nor / Per      2.936 0.944 Inf    1   3.349  0.0023
##
## Syntactic.Condition = Gender:
##  contrast  odds.ratio    SE  df null z.ratio p.value
##  L3 / Nor       0.966 0.273 Inf    1  -0.123  0.9917
##  L3 / Per       0.484 0.172 Inf    1  -2.044  0.1019
##  Nor / Per      0.501 0.176 Inf    1  -1.970  0.1197
##
## Syntactic.Condition = SV:
##  contrast  odds.ratio    SE  df null z.ratio p.value
##  L3 / Nor       0.992 0.265 Inf    1  -0.029  0.9995
##  L3 / Per       0.355 0.121 Inf    1  -3.048  0.0065
##  Nor / Per      0.358 0.120 Inf    1  -3.060  0.0063
##
## P value adjustment: tukey method for comparing a family of 3 estimates
## Tests are performed on the log odds ratio scale
```

Appendix 10.3 - Formula: Accuracy ~ 1 + age * Group + (1 | Participant.Public.ID) + (1 |sentences)

	Accuracy		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.29	0.03 – 2.90	0.291
age	1.23	1.02 – 1.48	**0.030**
Group [Nor]	150.24	4.82 – 4684.02	**0.004**
Group [Per]	1.41	0.07 – 27.02	0.818
age * Group [Nor]	0.67	0.51 – 0.88	**0.004**
age * Group [Per]	0.98	0.78 – 1.22	0.848
Random Effects			
σ^2	3.29		
$\tau_{00\ Participant.Public.ID}$	0.45		
$\tau_{00\ sentences}$	0.70		
ICC	0.26		
$N_{Participant.Public.ID}$	81		
$N_{sentences}$	60		
Observations	4768		
Marginal R^2 / Conditional R^2	0.039 / 0.289		

Appendix 10.4 - Formula: Accuracy ~ 1 + English.Proficiency * Group + (1 | Participant.Public.ID) + (1 |sentences)

	Accuracy		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.30	0.09 – 1.08	0.065
English Proficiency	1.13	1.06 – 1.20	**<0.001**
Group [Nor]	2.48	0.44 – 13.96	0.303
Group [Per]	1.35	0.13 – 14.17	0.803
English Proficiency * Group [Nor]	0.96	0.88 – 1.04	0.289
English Proficiency * Group [Per]	1.02	0.91 – 1.15	0.722
Random Effects			
σ^2	3.29		
$\tau_{00\ Participant.Public.ID}$	0.38		
$\tau_{00\ sentences}$	0.72		
ICC	0.25		
$N_{Participant.Public.ID}$	82		
$N_{sentences}$	60		
Observations	4828		
Marginal R^2 / Conditional R^2	0.051 / 0.289		

Appendix 10.1 - Formula: Grammatical Acceptance ~ Syntactic.Condition * Group + age + (1 | Participant.Public.ID) + (1 | sentences)

	Acceptance		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	29.96	4.56 – 196.94	**<0.001**
Syntactic Condition [AdvV]	0.55	0.22 – 1.35	0.191
Syntactic Condition [Definiteness]	1.77	0.68 – 4.60	0.244
Syntactic Condition [Gender]	1.27	0.50 – 3.24	0.613
Syntactic Condition [SV]	0.24	0.10 – 0.58	**0.001**
Group [Nor]	0.65	0.30 – 1.41	0.280
Group [Per]	3.39	1.25 – 9.22	**0.017**
age	0.89	0.78 – 1.03	0.113
Syntactic Condition [AdvV] * Group [Nor]	1.42	0.64 – 3.14	0.390
Syntactic Condition [Definiteness] * Group [Nor]	2.00	0.77 – 5.16	0.154
Syntactic Condition [Gender] * Group [Nor]	2.94	1.16 – 7.44	**0.023**
Syntactic Condition [SV] * Group [Nor]	1.46	0.68 – 3.13	0.328
Syntactic Condition [AdvV] * Group [Per]	1.31	0.47 – 3.62	0.608
Syntactic Condition [Definiteness] * Group [Per]	0.82	0.25 – 2.67	0.748
Syntactic Condition [Gender] * Group [Per]	2.37	0.64 – 8.81	0.196
Syntactic Condition [SV] * Group [Per]	1.23	0.48 – 3.19	0.663

Random Effects

σ^2	3.29
$\tau_{00\ Participant.Public.ID}$	0.90
$\tau_{00\ sentences}$	0.36
ICC	0.28
$N_{\ Participant.Public.ID}$	80
$N_{\ sentences}$	30
Observations	2364
Marginal R^2 / Conditional R^2	0.172 / 0.401

Appendix 10.2 - Post-hoc pairwise comparison of Appendix 14.4

```
## $contrasts
## Syntactic.Condition = AdjN:
##  contrast  odds.ratio    SE  df null z.ratio p.value
##  L3 / Nor       1.424 0.531 Inf    1   0.949  0.6095
##  L3 / Per       0.311 0.153 Inf    1  -2.377  0.0460
##  Nor / Per      0.218 0.104 Inf    1  -3.179  0.0042
##
## Syntactic.Condition = AdvV:
##  contrast  odds.ratio    SE  df null z.ratio p.value
##  L3 / Nor       1.133 0.398 Inf    1   0.356  0.9327
##  L3 / Per       0.254 0.116 Inf    1  -3.002  0.0076
##  Nor / Per      0.224 0.101 Inf    1  -3.330  0.0025
##
## Syntactic.Condition = Definiteness:
##  contrast  odds.ratio    SE  df null z.ratio p.value
##  L3 / Nor       0.892 0.395 Inf    1  -0.257  0.9641
##  L3 / Per       0.364 0.203 Inf    1  -1.810  0.1663
##  Nor / Per      0.407 0.228 Inf    1  -1.602  0.2449
##
## Syntactic.Condition = Gender:
##  contrast  odds.ratio    SE  df null z.ratio p.value
##  L3 / Nor       0.559 0.241 Inf    1  -1.350  0.3677
##  L3 / Per       0.133 0.084 Inf    1  -3.190  0.0041
##  Nor / Per      0.237 0.154 Inf    1  -2.219  0.0680
##
## Syntactic.Condition = SV:
##  contrast  odds.ratio    SE  df null z.ratio p.value
##  L3 / Nor       0.989 0.331 Inf    1  -0.032  0.9994
##  L3 / Per       0.248 0.104 Inf    1  -3.325  0.0025
##  Nor / Per      0.250 0.104 Inf    1  -3.337  0.0024
```

Appendix 10.3 - Formula: Ungrammatical Acceptance ~ Syntactic.Condition * Group + age + (1 | Participant.Public.ID) + (1 | sentences)

	Acceptance		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	19.96	1.90 – 209.45	**0.013**
Syntactic Condition [AdvV]	2.89	1.51 – 5.50	**0.001**
Syntactic Condition [Definiteness]	1.27	0.66 – 2.45	0.474
Syntactic Condition [Gender]	1.12	0.58 – 2.17	0.738
Syntactic Condition [SV]	0.84	0.43 – 1.66	0.624
Group [Nor]	0.51	0.20 – 1.34	0.172
Group [Per]	0.99	0.31 – 3.09	0.982
age	0.70	0.58 – 0.84	**<0.001**
Syntactic Condition [AdvV] * Group [Nor]	1.16	0.52 – 2.61	0.712
Syntactic Condition [Definiteness] * Group [Nor]	2.56	1.13 – 5.79	**0.025**
Syntactic Condition [Gender] * Group [Nor]	2.44	1.07 – 5.57	**0.034**
Syntactic Condition [SV] * Group [Nor]	1.88	0.80 – 4.40	0.147
Syntactic Condition [AdvV] * Group [Per]	1.17	0.48 – 2.88	0.729
Syntactic Condition [Definiteness] * Group [Per]	19.58	7.76 – 49.39	**<0.001**
Syntactic Condition [Gender] * Group [Per]	1.03	0.40 – 2.66	0.948

Syntactic Condition [SV] * Group [Per]	0.30	0.10 – 0.91	**0.034**

Random Effects

σ^2	3.29
$\tau_{00\ Participant.Public.ID}$	1.97
$\tau_{00\ sentences}$	0.09
ICC	0.38
$N_{\ Participant.Public.ID}$	80
$N_{\ sentences}$	30
Observations	2359
Marginal R^2 / Conditional R^2	0.193 / 0.504

Appendix 10.4 - Post-hoc pairwise comparison of Appendix 14.6

```
## $contrasts
## Syntactic.Condition = AdjN:
##  contrast  odds.ratio     SE  df null z.ratio p.value
##  L3 / Nor      1.5993 0.7931 Inf    1   0.947  0.6106
##  L3 / Per      1.0773 0.6442 Inf    1   0.125  0.9915
##  Nor / Per     0.6736 0.4012 Inf    1  -0.663  0.7848
##
## Syntactic.Condition = AdvV:
##  contrast  odds.ratio     SE  df null z.ratio p.value
##  L3 / Nor      1.5049 0.7057 Inf    1   0.872  0.6582
##  L3 / Per      0.9185 0.5200 Inf    1  -0.150  0.9876
##  Nor / Per     0.6103 0.3408 Inf    1  -0.884  0.6503
##
## Syntactic.Condition = Definiteness:
##  contrast  odds.ratio     SE  df null z.ratio p.value
##  L3 / Nor      0.6643 0.3160 Inf    1  -0.860  0.6657
##  L3 / Per      0.0546 0.0314 Inf    1  -5.058  <.0001
##  Nor / Per     0.0822 0.0460 Inf    1  -4.467  <.0001
##
## Syntactic.Condition = Gender:
##  contrast  odds.ratio     SE  df null z.ratio p.value
##  L3 / Nor      0.6870 0.3286 Inf    1  -0.785  0.7124
##  L3 / Per      1.0434 0.6178 Inf    1   0.072  0.9972
##  Nor / Per     1.5187 0.8770 Inf    1   0.724  0.7496
##
## Syntactic.Condition = SV:
##  contrast  odds.ratio     SE  df null z.ratio p.value
##  L3 / Nor      0.9855 0.4874 Inf    1  -0.030  0.9995
##  L3 / Per      3.6206 2.4192 Inf    1   1.926  0.1315
##  Nor / Per     3.6739 2.4266 Inf    1   1.970  0.1196
##
## P value adjustment: tukey method for comparing a family of 3 estimates
## Tests are performed on the log odds ratio scale
```