Does the Amount of English Exposure Affect the Size of the Grammaticality Illusion?

Patty Ernst

S4196724

Research Master's Thesis

Research Master in Language and Communication

Radboud University Nijmegen

Supervisors: dr. S. L. Frank & dr. R. Cozijn

July 2017

LIST OF CONTENTS

Abstract

This study investigates the influence of the amount of exposure to English on the missing-VP2 effect, a grammaticality illusion. It was expected that the size of the effect gets bigger when the amount of exposure to English increases. The missing-VP2 effect occurs when in double centre-embedded relative clauses, the second (middle) verb phrase is left out. Sentences without this VP2 are not grammatically correct, but are read faster or understood better than the sentence in which the second verb phrase is present, the grammatically correct sentence. Based on study background and English exposure, participants were placed in either a Dutch or English language group. Eye Tracking was used to measure reading times on double centred-embedded relative clauses, which were shown in four conditions: (1) semantically biased - grammatical, (2) semantically biased – ungrammatical, (3) semantically neutral – grammatical, and (4) semantically neutral - ungrammatical. The task was to read the sentences and occasionally answer a content question. The results showed that the grammaticality illusion only occurred in semantically neutral sentences. There was no significant difference between the two language groups, and no clear effect of exposure. However, a trend was visible in the data in which a higher amount of English exposure could increase the size of the grammaticality effect with a bigger sample size.

Keywords: Amount of exposure; Eye-tracking; Grammaticality Illusion; Missing-VP2 effect; Reading comprehension

1. Introduction

When thinking about illusions, one would probably imagine a magician pulling a rabbit out of a hat, or cutting a girl in half. However, illusions do not only occur in the world of magic. For instance, there are visual illusions in the famous paintings of M.C. Escher with endless staircases and two hands drawing each other perpetually. There are also illusions in the domain of language, for instance, when grammatically incorrect sentences are perceived as correct.

What is meant by the term 'grammatically incorrect'? These are sentences that are not in line with the syntax of a specific language. One can experience an illusion of grammaticality in a situation where a grammatically incorrect sentence is perceived as grammatical in terms of comprehensibility judgements or when the grammatically incorrect sentences are read faster than the grammatically correct sentences. This can occur, for example, when the subject-verb agreement is incorrect: 'The time for fun and games are over' and 'The key to the cabinets are lost' (Bock & Miller, 1991; Vigliocco & Nicol, 1998). In both cases the verb 'are' should be replaced by 'is', because the subjects are singular ('the time' and 'the key'), but in most cases this is not noticed. In this paper, the phenomenon is investigated in double centre-embedded relative clauses. Consider the following example:

(1) 'The song that the mother who the children scared in the living room thoroughly enjoyed singing was about a lost love.'

As you can see in (1) three sentences are combined: (1) 'The song was about a lost love', (2) 'the mother thoroughly enjoyed singing (the song)', and (3) 'the children scared (the mother) in the living room'. By embedding these three sentences into one, that final sentence becomes long, complex, and difficult to read. However, it still is a sentence that is possible following English grammar rules. Several studies have already shown that in English and

French a grammaticality illusion, also called the *missing-VP2 effect*, occurs when the second verb phrase (VP2) is left out. Consider the next example:

(2) 'The song that the mother who the children scared in the living was about a lost love.'

The verb phrase 'thoroughly enjoyed singing' is left out in sentence (2) which makes it grammatically incorrect. However, these types of sentences in English are judged as more comprehensible than the grammatical counterpart by English natives and Dutch natives (Frank & Ernst, submitted; Gibson & Thomas, 1999), and resulted in lower reading times in self-paced reading and eye-tracking tasks in German, Dutch and English participants (Christiansen & MacDonald, 2009; Frank, Trompenaars, & Vasishth, 2016; Vasishth, Suckow, Lewis, & Kern, 2010).

From previous research it can be concluded that the grammaticality illusion occurs with speakers from different languages. Recent studies by Frank and colleagues (Frank et al., 2016; Frank & Ernst, submitted) showed that Dutch and German participants with high English proficiency, reading English double centre-embedded relative clauses, also showed a grammaticality illusion as indicated by the reading times of the missing-VP2 condition (the condition in which the second verb phrase is left out). The reading times of the third verb phrase in the missing-VP2 condition were lower than in the grammatical condition. However, when these participants read the stimuli in their native language (Dutch or German), no illusion occurred, which suggests that the high proficiency level of participants is involved in this process, and that the amount of exposure to the second language possibly influences the grammaticality illusion.

This thesis will investigate the influence of the amount of exposure to English as a second language on the size of the grammaticality illusion effect. It will give a broad overview of the literature about this topic, discussing both offline and online measurements of

grammaticality illusions in double-centre embedded relative clauses, and describe the eye-tracking experiment that has been conducted. The thesis concludes with a discussion of the results.

2. Theoretical background

2.1 The grammaticality illusion

A considerable number of studies has already been conducted concerning the grammaticality illusion. One of the first experimental studies that investigated this illusion was a study by Gibson and Thomas (1999). They collected simple comprehensibility judgement scores on a Likert scale. In their stimuli, they left each of the three verb phrases out, which created four sentence conditions: three ungrammatical ones in turn omitting one of the verb phrases, and one grammatical one with all the verb phrases present. The following examples show the four conditions:

(3) 'The ancient manuscript that the grad student who the new card catalog had confused a great deal was studying in the library was missing a page.'

a. <u>Missing VP1:</u> The ancient manuscript that the grad student who the new card catalog was studying in the library was missing a page.

b. <u>Missing VP2:</u> The ancient manuscript that the grad student was who the new card catalog had confused a great deal was missing a page.

c. <u>Missing VP3:</u> The ancient manuscript that the grad student who the new card catalog had confused a great deal was studying in the library.

An important aspect about this study is that the target sentences had a complex and fixed structure and had strong selectional restrictions on the verbs. The second noun phrase and verb phrase were semantically biased, for instance, the verb 'to shred' in sentence (3) is one of the few verbs that can be used with the noun 'paper shredder'. However, when using 'the mother' or 'the student' almost every verb can follow.

The hypotheses tested by Gibson and Thomas (1999) were based on several existing theories about *memory limitations* and *structural forgetting*. First of all, the *Disappearing Syntactic Nodes Hypothesis* by Frazier (as cited in Gibson and Thomas, 1999) predicts that the ungrammatical conditions have lower scores than the grammatical conditions. The *Least Recent Nodes Hypothesis* (Gibson, 1991), the second theory that was tested, predicts that the first noun phrase will be forgotten, which in turn leads to the unexpected encounter of the third (last) verb phrase that goes with it. Thus, conversely, the omission of the third verb phrase would not be noticed and the sentence would have an equal or higher score in comparison to the grammatical condition, in which all verb phrases are present. The third theory is the *Syntactic Prediction Locality Theory* (SPLT; Gibson, 1998), which predicts that the omission of the second verb phrase will cause a grammaticality illusion. It underlines the influence of *locality* and of integration costs (i.e., the number of referents between the verb and the dependent) and memory costs (i.e., the number of heads that need to be remembered) on sentence comprehension. The term locality indicates how 'local' the integration is. The cost of integrating two elements in a sentence depends on the distance between them. The longer the distance between a noun phrase and a verb phrase, the more effort it takes to integrate them. The SPLT indicates the omission of the second verb as the cause of the effect by focussing on the overall integration costs at the verb phrases. The integration cost is the highest at the second verb phrase, so leaving that component out would lower the integration costs the most, resulting in a grammaticality illusion.

The results of the experiment by Gibson and Thomas (1999) showed that the sentences in which the second (middle) verb phrase was left out were scored as equally comprehensible as the grammatically correct sentences. The sentences in which the first or third verb phrase was left out had significantly lower comprehensibility scores. These results are in line with the SPLT (Gibson, 1998). It was not the case that the sentence in which the second verb

phrase was left out scored better on the comprehensibility measure. The authors argued that because the participants had the opportunity to look back and read the sentence again, the difference between comprehensibility scores of the grammatically correct condition and the 2nd verb phrase condition was not significant.

An aspect that remained unclear was whether the semantic bias that was present in the stimuli could have had an influence on the results. This was further investigated in a study by Christiansen and MacDonald (2009), who conducted four experiments in which they explored a usage-based perspective on recursive sentence processing, also in double centre-embedded relative clauses. Only the third and fourth experiment are relevant for this thesis, so they are discussed here.

In the third and fourth experiment of Christiansen and Macdonald (2009), participants had to determine for every word if the sentence up to that point made sense or not. When the sentence was judged as ungrammatical, a question was asked about how 'good' the sentence was on a 7-point Likert-scale. In the third experiment, the semantically biased stimuli from Gibson and Thomas (1999) were tested. The results showed that significantly more sentences with three verb phrases (grammatical ones) were rejected and had lower grammaticality ratings than sentences with two verb phrases. Christiansen and Macdonald (2009) state that there could have been a confound in the stimuli of Gibson and Thomas (1999) between grammatical condition and sentence length. The stimuli were not controlled for length and, therefore, could have been biased towards the 2VP sentences as they were shorter. They also state that the presence of a semantic bias in the stimuli of Gibson and Thomas could be a confound. Therefore, Christiansen and MacDonald (2009) conducted a fourth experiment with semantically neutral sentences. However, the results showed the same pattern as in the third experiment: The sentences with 3VPs were rejected more often and had lower grammaticality ratings than the 2VP sentences.

The findings from Christiansen and Macdonald (2009) support the notion that also comes forward in the study by Gibson and Thomas (1999). When native English speakers read English sentences, there is a grammaticality illusion: The grammatically correct sentences (sentence with 3VPs) are rejected more often and have lower grammaticality ratings than the grammatically incorrect sentences, in which the second verb phrase is left out. The study also shows that there is no difference between the semantically neutral and the semantically biased sentences.

These findings are also replicated in French. A self-paced-reading study by Gimenes, Rigalleau, and Gaonac'h (2008) investigated whether the missing-VP2 effect also occurred in French with native French speaking participants. The study by Gibson and Thomas (1999) showed that this illusion only occurred when the second verb phrase was missing, so Gimenes et al. only used two conditions: one with the second verb phrase missing and one in which all three verb phrases were present. The stimuli were shown to the participants in a self-paced reading experiment. The sentences contained verbs with a semantic bias. After reading a sentence, a question was asked that had to be answered orally. For every target sentence, two questions were constructed, of which one was presented depending on the condition of the sentence: One asked what the second noun phrase (NP2) did, and the other asked what the third noun phrase (NP3) did. The participants also had to indicate the complexity of the sentence after each trial on a 5-point scale.

The results showed that the missing VP2-effect also exists in French. The condition in which the second verb phrase was missing was easier to process than the grammatical condition in which all three verb phrases were present. However, this was only the conclusion of the comprehension questions and the complexity judgements, but not from the reading times from the self-paced-reading task. The reading times of the third VP in the missing VP2-condition were longer than in the grammatical condition. The authors state that these results

reflect a position effect, meaning that later positions are read faster than earlier positions. Another explanation they give is that the second verb phrase is often not 'forgotten' because of the frequent questions about that specific component in other trials.

The above mentioned studies (Christiansen & MacDonald, 2009; Gibson & Thomas, 1999; Gimenes et al., 2009) claim that the explanation of this phenomenon lies in the working memory constraints of the participants, which cause the structural forgetting of the second subject in the sentence. It is argued that working memory does not have enough capacity to store three embedded structures, which causes the illusion of grammaticality, when the second verb phrase is omitted. To investigate whether working memory constraints really are a possible cause, instead of other linguistic constraints, multiple cross-linguistic studies were conducted. These will be discussed in the next chapter.

2.2 Cross-linguistic views: detecting the illusion via online methods

A study by Vasishth, Suckow, Lewis, and Kern (2010), showed that the explanation that working memory is the cause of the grammaticality illusion effect may not be completely valid. They conducted six experiments to see whether the grammaticality illusion could be detected by online methods, such as eye-tracking and self-paced reading, and not only by offline methods, such as comprehensibility judgements. The study by Gimenes et al. (2009) only detected the grammaticality illusion in the sentence ratings and not in the reading times of a self-paced reading task.

German and English participants were tested in their native languages to see whether in both languages the grammaticality illusion occurred in self-paced-reading tasks and eye-tracking tasks. The languages German and English were chosen because of the different language structures: In the sentences that were investigated, German has a subject-object-verb (SOV) structure whereas English has a subject-verb-object (SVO) structure. This also means that English is a *head initial* language, while German is a *head final* language. In a head final

language, the 'head' of the sentence (the constituent that determines the category of that utterance, which is mostly a verb) follows after most other parts of the sentence and thus is in sentence final position (Polinsky, 2012).

The goal of the study was to detect a grammaticality illusion in a language with a different word order than English, through measuring reading times. Six experiments were conducted: The first two were a self-paced-reading study and an eye-tracking study with English participants, the third and fourth experiment were similar but in German with native German speakers. In the German stimuli, commas were added in order to strengthen the ecological validity as these are compulsory in German relative clauses. To test if this also made a difference in English, another self-paced-reading task was done, with commas added to the English stimuli. The last experiment consisted of English sentences in which the relative pronoun was always 'who', which was not the case in the previous English experiments. The sentences used in all the experiments had the same structure as those of Gibson and Thomas (1999), however, without a semantic bias, and there were two conditions: a grammatical condition with all three verb phrases present, and an ungrammatical condition in which the second verb phrase was left out. After each sentence, a yes/no comprehension question was asked.

This study was based on the *Dependency Locality Theory* (DLT, Gibson, 2000), which is a more elaborated and adapted version of the SPLT (Gibson, 1998). The DLT differs from the SPLT in some definitions, but it still has the same two key components: integration and locality. The integration component was extended with a storage component, which consists of the 'storage of the structure built thus far', (Gibson, 2000, p. 102) while the integration consists of the 'integration of the current word into the structure'(Gibson, 2000, p. 102). They both predict that the omission of the second verb phrase in double-centre embedded relative clauses causes the grammaticality illusion.

The results showed that there was no grammaticality illusion effect in German, neither in the self-paced-reading study not the eye-tracking study. Contrarily, in English, the illusion was found in both tasks, also in the fifth experiment in which commas were added. In English, the participants read the third VP of the sentence in which the second verb phrase was missing faster than the grammatical counterpart with three verb phrases. In German, the opposite was the case: The ungrammatical condition had higher reading times than the grammatical counterpart in both the self-paced-reading and the eye-tracking experiment. This means that no grammaticality illusion was found in German, whereas it was found in English.

This discrepancy in the results of Vasishth et al. (2010) can be explained by the word-order differences between English and German. The head final structure of German may cause verbs to be more salient in the verbal working memory of German speakers than of English speakers. On the other hand, the structure of English may make the verbs less salient in the verbal working memory of English speakers. If this is the cause of the difference between the English and German results, the German participants are probably more skilled in remembering verbs, which makes their working memory robust for the structural forgetting of one of the subjects in the sentence, and thus insensitive to the grammaticality illusion, resulting in ungrammatical sentences to be read slower than grammatical sentences.

 Based on the findings reported by Vasishth et al. (2010), Haüssler and Bader (2015) propose an interference account to explain the missing-VP effect by means of the *Discrimination Hypothesis*. This view states that the integration of an NP with the associated VP becomes harder when there is an intervening phrase, that contains an alternative for the VP. Specifically, in a double centre-embedded relative clause, there are two possible noun phrases that the verb phrase can be attached to, which are hard to distinguish from each other. Due to primacy effects, it is assumed that the 'choice' will be more often for the first noun phrase, instead of the second, which could cause a grammaticality illusion, the missing-VP2

effect. Primacy and recency effects are part of serial position effects, as defined by Murdock (1962). He denotes that information that comes first (primacy) or last (recency) is remembered the best.

The interference account was tested by conducting two experiments. Häusler and Bader (2015) proposed to make a distinction between the results found on the missing-VP effect and the choice of method, i.e. offline (grammaticality judgements) and online (self-paced reading and eye tracking). They also tested whether German natives could have a grammaticality illusion in German, and not only a missing-VP2 effect, but also a missing-VP1 effect, by looking at the type of clause. The previous experiments all looked at relative clauses within a main clause, but this experiment investigated relative clauses within a complement clause. An example of such a sentence structure is: "I have heard that the intern who caused the system crash which occupied the engineer for several hours disappeared since noon." Following the discrimination hypothesis, the two options of integration of the second verb phrase (to the NP1 and NP2) should both be possible. By testing these two different clause types in combination with missing VP1 or VP2, the influence of primacy effects on this effect can be determined, which would support the interference account.

The first experiment consisted of a questionnaire measuring binary grammaticality judgements (grammatical or ungrammatical) in six different sentence conditions: The sentence could be either a main clause or an embedded clause, and it could be either complete, missing the VP1, or missing the VP2. There was no time limit in reading the sentences. The results showed that even when participants have unlimited time to read the sentences, there still is a grammaticality illusion effect. In general, the two missing-VP conditions were substantially often rated as grammatical. A second conclusion is that both in the main clause and in the embedded clauses a missing-VP2 effect was present, while the missing-VP1 effect was only found in the embedded clauses.

The second experiment, a self-paced reading task, investigated whether the missing-VP2 effect could also be found in online reading times. In order to do so, they, in addition, manipulated grammatical number in the sentences (subject-verb agreement). There were four sentence conditions: The VP2 was either missing or not, and the second NP (belonging to the second VP) was either plural or singular. When second VP and verb are plural and the second VP is left out, there will be a subject-verb agreement error when trying to process the sentence, since the VP1 and NP1 were singular in all conditions. The results showed that the reading times for the third, final verb were shorter in the missing-VP2 condition than in the grammatical counterparts. Haüssler and Bader (2015) conclude that the effect not only takes place in grammaticality judgements, but also when the participants' task is just to read for comprehension. There was no effect of number manipulation, which is attributed to the absence of a grammaticality judgement in the second experiment. Based on their findings, Haüssler and Bader (2015) proposed the interference account (explained above) to explain the missing-VP2 effect.

Another explanation that could account for the occurrence of the missing-VP2 effect is proposed by Frank et al. (2016). This cross-linguistic study tried to determine whether the grammaticality illusion effect is language specific or is caused by working memory limitations, as working memory limitations cannot be language specific. A self-paced-reading task was conducted with Dutch and German participants to measure reading times. Dutch also has a head final structure, just like German. The first experiment tested Dutch participants with Dutch stimuli, and the second and third experiment tested German and Dutch participants in English, their second language. The stimuli, again centre-embedded relative clauses, consisted of one grammatical and one ungrammatical condition: One in which all three verb phrases were present, and one in which the second verb phrase was omitted. There

was no semantic bias in the verbs of the stimuli, unlike the stimuli of Gibson and Thomas (1999).

The results showed that there was no grammaticality illusion: The grammatical condition was read faster than the ungrammatical condition in the Dutch sentences, which resembled the findings in German by Vasishth et al. (2010). However, an important finding was that the reading times of the English sentences of the German and Dutch participants differed between the two conditions, supporting the conclusion that there was a grammaticality illusion effect when reading a second language. This means that the Dutch and German participants have the same reading behaviour as English natives when reading English missing-VP2 sentences. They read those sentences faster than the grammatical counterparts with all three verb phrases present. This finding indicates that this specific phenomenon might be language specific. The occurrence cannot be explained by working memory constraints, because Dutch and German natives do have a missing-VP2 effect when reading English, but not in their native language. To account for the results, the authors propose a language statistics account, which states that exposure to a language causes sensitivity for that language's specific statistics and word-order patterns. The higher the exposure is, the stronger the sensitivity is, and in case of the missing-VP2-effect, the stronger the missing-VP2-effect is.

So far, three possible explanations for the occurrence of the grammaticality illusion have been proposed: the structural forgetting account (Gibson & Thomas, 1999), the interference account (Häussler & Bader, 2015), and the language statistics account (Frank et al., 2016). Frank et al. (2016) conclude that the three accounts that have been proposed so far do not fit the results. Frank and Ernst (submitted) suggest an account that combines a statistical and a working-memory approach.

The study by Frank and Ernst (submitted) was a follow-up of the study by Frank et al. (2016). Three experiments were conducted to further analyse the occurrence of the grammaticality illusion in Dutch and English. The first experiment replicated the study by Gibson and Thomas (1999) in Dutch, by translating the sentences and testing them with comprehensibility and acceptability ratings to see whether this rating method would reveal a grammaticality illusion in Dutch with that specific method. The sentences occurred in four conditions: missing the first, the second, or the third VP, and a condition in which all three VPs are present. The participants rated the sentences via an online questionnaire, indicating per sentence on a 7-point scale how comprehensible and how acceptable it was. The results showed that the grammatical condition was rated higher on comprehensibility and acceptability, which means that no grammaticality illusion was found in the first experiment, which corresponds to the findings of Frank et al. (2016).

The goal of the second experiment was to replicate the findings of Gibson and Thomas (1999), by using the method of the first experiment, but converting and translating it into English, and testing it with English native speakers. The results showed that the missing-VP2 sentences had higher ratings than the grammatical sentences, which indicates a grammaticality illusion. The results were therefore consistent with the results found by Gibson and Thomas.

Finally, the third experiment combined the two former experiments in a within-subjects design. Dutch native speakers with high English proficiency were exposed to both the English and the Dutch stimuli, in two separate language blocks. The stimuli of experiments 1 and 2 were used, and the method was again the same. One of the differences with the first two experiments was that the third experiment took place in a controlled lab setting, instead of online. The task was again to rate the sentences on comprehensibility and acceptability. In addition, the participants were asked to take part in an English reading

proficiency test (the Vernon-Warden reading test; Hedderly, 1996), a regular English

proficiency test (LexTALE; Lemhöfer & Broersma, 2012), and a language background

questionnaire. The results of the rating task in the third experiment showed that there was no

evidence of a grammaticality illusion in the Dutch sentences, but that there was one in the

English sentences. Overall, the effect does appear in English, but not in Dutch, even within

the same participant group of Dutch natives with English as L2 and by using grammaticality

judgements as a measure.

The authors claim that the results do not fit any of the three already existing accounts,

and propose an account that combines some of the already existing accounts: a hybrid

statistical and working-memory account. This account states that both primacy and recency

effects are the cause of the 'forgetting' of the second NP. Additionally, the fact that sentences

with three consecutive VPs do not occur frequently in English may cause them to be

perceived as more grammatically incorrect than the more frequent sentences with two

consecutive VPs. In Dutch and German, sentences with three consecutive VPs are more

frequent and are, therefore, more acceptable by native speakers.

To briefly summarize the findings from the above two chapters: The grammaticality

illusion was found in English and French when leaving the second verb phrase out in self-

paced reading scores, eye tracking measures, and comprehensibility judgements. The illusion

was not found in Dutch and German, using the same methods, however when German and

Dutch participants read English stimuli, the grammaticality illusion did appear in missing-

VP2 sentences. Although, Haüssler and Bader (2015) did find a grammaticality illusion in

German sentences when using a slightly different sentence structure. An explanation, would

be that in languages with a certain structure, a limitation in working memory causes structural

forgetting processing, irrespective of whether that language is the native language of the

reader or not. These findings support the notion that the differences between the findings of

English and French on the one hand and Dutch and German on the other are caused by differences in the internal structure of the language, i.e., English and French are head initial languages, whereas Dutch and German are head final languages. The language statistics have an influence on working-memory, through primacy and recency effects.

The fact that there is no cross-linguistic influence when reading in a second language with this phenomenon, is an unexpected finding. Multiple studies in the past have shown that when learning a second language, there is cross-linguistic influence from the L1 to the L2 and vice versa (Cenoz, 2001; Saville-Troike, 2013). It is also reasonable to assume that one reads a sentence with the same structure similarly in their second language as in the native language. That is apparently not the case with these particular complex double centre-embedded relative clauses. These findings could perhaps be caused by differences in proficiency and exposure. For example in Frank et al. (2015) English proficiency tests were conducted to measure a certain level of proficiency, but there were no analyses of the relation between proficiency or exposure and the size of the missing-VP2 effect. This could, however, provide insights in the nature of this phenomenon. Does the size of the effect differ when the exposure to English, in both reading and listening, increases or decreases? Is it the case that more exposure to English results in a larger effect than less exposure? The answers to these questions may elucidate whether the grammaticality illusion really is language specific or not.

2.3 L2 proficiency and exposure

Frank et al. (2016) and Frank and Ernst (submitted) used proficiency tests to ensure a high level of proficiency with participants, but they did not link proficiency or exposure to the occurrence of the grammaticality illusion. A question could be whether a high proficiency or high exposure to the language would make a difference and whether the strength of the phenomenon would gradually differ or not. As it has been assumed that this phenomenon relates to language structures and language statistics (Frank et al. 2016; Frank & Ernst,

submitted; Vasishth et al., 2010) it could be the case that the amount of exposure to a specific language has an influence. However, it is hard to make a clear distinction between proficiency and exposure, as they are also both hard to distinguish in terms of measuring. Amount of exposure is often defined as the length and the frequency someone has been in contact with a language (Saville-Troike, 2013, p. 112), while proficiency is about how 'skilled' someone is in a (second) language. Multiple studies show that there is a correlation between proficiency and amount of exposure, maybe even more than between starting age and proficiency (Ojima, Matsuba-Kurita, Nakamura, Hoshino, & Hagiwara, 2011). The higher the amount of exposure to a (second) language, the higher the proficiency scores are. The conclusion of those studies is that the amount of exposure has a great influence on ultimate proficiency.

Previous research (Frank et al., 2015; Frank & Ernst, submitted) has shown that highly proficient second language learners of English were also sensitive to a grammaticality illusion when reading English, as were native English speakers. However, it is unclear whether the amount of exposure was the underlying cause or not. It seems to be the case that participants had been exposed enough or were proficient enough to be sensitive to the language statistics of English, and to use those when reading English sentences like English natives do. The present study investigates whether the grammaticality illusion effect in English is affected by exposure to English, by testing whether the grammaticality illusion effects gets bigger when the amount of English exposure gets higher.

2.4 Present study

The current study aims to further explore the nature of the grammaticality illusion by also looking at and controlling for the exposure to English. By looking at two different participant groups, based on their exposure to English, it is possible to investigate whether the effect will increase or decrease when the participants are more exposed to English. It is possible that the increase of English exposure as an second language, causes more sensitivity

to language structures of English and therefore causes a bigger grammaticality illusion. One group of participants that was mostly confronted with Dutch in their study program was compared with a second group of participants that was mostly confronted with English in their study program.

The goal of the experiment was to answer the following research question: 'What is the role of English exposure of Dutch natives in the size of the grammaticality illusion effect, in reading double centre-embedded relative clauses?' We have three main expectations. The first hypothesis is that there will be an overall grammaticality illusion, because the sentences are read in English. It is expected that in all groups and conditions it is the case that the third verb phrase in the ungrammatical missing-VP2 condition is read faster than in the grammatical condition. The second hypothesis is that in the English language group the size of the grammaticality illusion effect will be bigger than in the Dutch language group, i.e. the difference between the reading times of the third verb phrase in the grammatical and ungrammatical sentences is expected to be bigger in the English language group than in the Dutch language group. It is also expected that study year (and thus amount of exposure in years) has a strengthening influence on the size of the effect. Consequently, the third hypothesis is that in the English group, the higher the amount of exposure, the larger and more robust the effect is, i.e. the bigger the difference will be between the ungrammatical and grammatical condition. In the Dutch group, there will be no influence of amount of exposure, because effectively, the amount of exposure to English does not get significantly higher over the years, so the effect should not differ between the different study years.

Finding a difference in the strength of the missing VP2-effect between the two groups would reveal something about the nature of the effect. If the strength of the effect is stronger in the English exposure group than in the Dutch exposure group, it can be assumed that the effect is not caused by working memory constraints. A larger exposure to English should have

no influence on working memory capacity, but it could have influence on the size of the grammaticality illusion. It is likely to say that if the effect would be caused by working memory constraints, the grammaticality illusion effect would not get bigger.

To get a better view of the participants' reading times, and to get more insight into the reading processes, eye tracking was used as a method. In previous studies, grammaticality judgements (offline) and self-paced-reading (online) were used. Grammaticality judgements are subjective and conscious decisions and in self-paced-reading, the reading process is interrupted by the pacing, which makes it useful to try a different method. Eye tracking is a frequently used method in reading comprehension research, because it can track the exact movements of the eyes. These eye movements reflect the cognitive processing that happens during reading (Rayner, 1998). This method produces the exact reading times of a sentence and its separate words and may also show which parts of a sentence readers find harder to process.

3. Method

3.1 Materials

There were 24 target sentences, 12 semantically biased and 12 semantically neutral sentences. The 12 semantically biased sentences were taken from the study by Gibson and Thomas (1999), and the 12 semantically neutral sentences originated from Frank et al. (2016). The target stimuli were adjusted for length, to exactly fit the screen in the experiment. By putting a line break in the sentences, it was made sure that the third verb phrase and one or two words before that would be completely on the second text line, and that the rest of the sentence would fit on the first text line, without going beyond the boundaries of what the eye tracker could measure. Some words from the first line were replaced by shorter synonyms or similar words, without changing the meaning of the sentence. The list of stimuli with both conditions and the line breaks can be found in Appendix A. The exact changes made to the

target sentences in comparison to the stimuli of previous studies can be found in Appendix B. These changes were only made to fit everything on the screen or to make the sentences more comprehensible to non-native speakers. Some difficult English terms were changed into similar and easier words to ensure that the all participants would know them. The verb in the third verb phrase was never at the end of the text line, to ensure that the participants would not read it only parafoveally, and extra adverbs were inserted, as the last words of a sentence are often processed parafoveally (Schotter, Angele & Rayner, 2012). The original semantically neutral stimuli did not have any adverbs in the verb phrases, only verbs, while in the semantically biased stimuli the verbs were always followed by some more information. To make the two types of sentences more similar, more adverbs were added to the verb phrases in the semantic neutral stimuli, if they did not already contain some.

The sentences used in the experiment were double centre-embedded relative clauses, with three consecutive noun phrases (NP1, NP2, and NP3), followed by three consecutive verb phrases (VP1, VP2, and VP3). The sentences were presented in two conditions: grammatically correct and grammatically incorrect. In grammatically incorrect sentences, the second verb phrase was omitted. This resulted in target sentences in four conditions, as shown in Table 1.

Table 1

Examples of the four sentence conditions: The columns show the Semantic Bias condition, while the rows show the Grammaticality condition.

	Semantically biased	**Semantically neutral**
Grammatically correct	[The book]$_{NP1}$ [that the student]$_{NP2}$ [who the new catalog]$_{NP3}$ [had confused a great deal]$_{VP1}$ [was studying in the library]$_{VP2}$ [was missing an important page.]$_{VP3}$	[The carpenter]$_{NP1}$ [who the craftsman]$_{NP2}$ [who the peasant]$_{NP3}$ [carried a long way]$_{VP1}$ [hurt on purpose]$_{VP2}$ [supervised the apprentice in the garden.]$_{VP3}$
Grammatically incorrect	[The book]$_{NP1}$ [that the student]$_{NP2}$ [who the new catalog]$_{NP3}$ [had confused a great deal]$_{VP1}$ [was missing an important page.]$_{VP3}$	[The carpenter]$_{NP1}$ [who the craftsman]$_{NP2}$ [who the peasant]$_{NP3}$ [carried a long way]$_{VP1}$ [supervised the apprentice in the garden.]$_{VP3}$

In order to mask the purpose of the study 96 filler sentences were added. The fillers were partly constructed from the fillers used by Koster (2016), and mostly from fillers used in Frank et al. (2016). To be sure the fillers were in correct English, they were checked by an English native speaker.

The font used for the sentences in the experiment was Calibri (size 18), and the text was always one tab away from the edge of the screen, both in front and behind the text. This was to make sure that the gaze of the participants would be within the limits of the eye tracker. For the instruction screens, Times New Roman (size 20) was used, to make a clear difference between the sentences and the instruction screens. This is also the reason why the questions were size 20 in Calibri. The questions screens consisted of the word 'question', under that the question would follow (centred) and underneath that 'yes' and 'no' with the

corresponding keys 'z' and 'm'. To make sure that participants would read attentively and to check whether they understood the sentences, 30 sentences, 6 target sentences and 24 filler sentences, were followed by a comprehension question

3.2 Design

The independent variables were Grammaticality with conditions grammatical and ungrammatical, Semantic Bias with conditions biased or neutral, Group with the conditions English and Dutch, and Amount of exposure in number of years. The grammatical sentence had three consecutive verb phrases, whereas in the ungrammatical condition the second (middle) verb phrase was left out. The variable Group was determined by looking at the language questionnaire. The participants were placed in one group or the other, when one language (English or Dutch) would be present for more than 70 % in the study program, counting both the language used in the lectures as the language used in the literature the students had to read.

The Amount of exposure in years to Dutch or English was measured by looking at their years of study and their study program. For example, being in the third bachelor year of the study English language and culture counted as three years of English exposure. When a participant only received English exposure in their master program and not in their bachelor program, only the years in the master program were counted. The years of exposure to Dutch were measured in the same way.

A mixed design was used. Every participant was exposed to all four conditions. The variables Grammaticality and Semantic Bias were within-subjects factors, while Group and Amount of exposure were between-subjects factors. Every participant saw every target sentence in only one condition to avoid that the participant would read the target sentences in both conditions, which may give away the goal of the study. The target sentences were counterbalanced by condition, so four lists were constructed. The lists were constructed

manually. Firstly, one list was created, in which the order of the fillers, the four conditions of target sentences, and the sentences with questions were semi-randomly distributed. This ensured that the fillers and target item conditions were distributed evenly over the list. The second list had the reversed order of the first one. The third list was the same as the first list, but with the Grammaticality condition mirrored: grammatical/ungrammatical sentences on list 1 were changed into ungrammatical/grammatical sentences on list 3. Similarly, the fourth list was a mirror version of list 2. The 30 comprehension questions were semi-randomized in such a way that there always was at least one trial without a question between two trials with questions. The participants were administered randomly to one of the four lists.

3.3 Instrumentation

The reading times of the third verb phrase, the first pass gaze durations, were used as a dependent variable. The first pass gaze duration is measured from the first time the third verb phrase is encountered, not having been skipped in an earlier pass, until it is left in a forward or backward direction. The reading times of the third verb phrase were analysed, because they were expected to show an effect of the grammaticality illusion in English. Previous studies also used the reading times of the third verb phrase to measure the presence of a grammaticality illusion (Haüssler & Bader, 2015; Frank et al., 2016; Vasishth et al., 2010). Generally, when a sentence is difficult to read or comprehend, the reading time is longer compared to the time it takes to read a sentence that is less difficult (Rayner, Chace, Slattery, & Ashby, 2006). To determine whether the participants were paying attention during the experiment, 30 yes/no comprehension questions were added.

3.4 Participants

To investigate whether there was a difference in the size that the grammaticality illusion occurs, two groups were tested: one group with little exposure to English and one group with more exposure to English. In total, 57 participants were tested in this study, of

which the data of 16 were excluded. The exclusions were caused by technical problems, such as a failure to calibrate (one participant), a system crash (one participant), and faulty measurements (five participants), and participant characteristics, such as dyslexia (one participant) and a study background that was not strictly English or Dutch oriented (eight participants). Of the remaining 41 participants, the English group consisted of 18 participants (13 females, age range 19-25 years, mean age 21.9) and the Dutch group of 23 participants (18 females, age range 20-26, mean age 22.6). The exposure to English was controlled for by looking at the study program the participants were in, and by looking at the answers given in the language questionnaire.

The participants were also placed in one of four groups based on their amount of exposure to the English or Dutch language. In the English group, nine participants had one year of English exposure, four had two years of exposure, four had three years of exposure, and one had four or more years of exposure. In the Dutch language group, one participant had one year of Dutch exposure, five had two years of exposure, five had three years of exposure, and 12 had four or more years of exposure. This overview is also shown in Table 2.

Table 2

Overview of the number of participants per language per exposure year group.

	Amount of Exposure			
Group	One year	Two years	Three years	Four years or more
English group	9	4	4	1
Dutch group	1	5	5	12

To be allowed to participate, the students had to be enrolled for or have just finished a bachelor and/or master program, and had to be Dutch native speakers. Participants were allowed for the Dutch group from the following bachelor or master programs: Dutch

Language and Culture, Dutch Law, Medicine, PABO (Primary School Teacher Education),

Sports and Movement, Language and Speech Pathology, Advanced Business Creation, and

the ALPO (Academic Primary School Teacher Education). For the English group, students

with one of the following study programs were allowed: English Language and Culture,

Environment and Society studies, Nutrition and Health, Psychology (English bachelor),

International Law, International Business Communication, Linguistics (English master),

Language and Communication (English master), International Business Administration,

Communication and Information Sciences (track IBC, language: English), Medical Biology

and Economics. The participants registered via the Radboud SONA-system and they were

granted either a gift card of 10 euros or one course credit for their participation. Ethical

approval was granted by the Ethics Assessment Committee of the Faculty of Humanities of

the Radboud University Nijmegen.

3.5 Procedure

Before taking part in the study, the participants read the study information and signed

the consent form. The study information document informed the participants that they were

going to read sentences that were cut off at a certain point and were put on the next text line.

The document stated that the goal of the study was to investigate the English reading

behaviour of Dutch natives with English as a second language. Next, the dominant eye of the

participants was determined, by doing a simple task, and that eye was then appointed as the

eye to be tracked by the eye tracker. Before the start of the eye tracking experiment, the

participants were given verbal instructions. They were instructed to read the sentences like

they would normally read a text, to blink as little as possible, and not to move during the

experiment. The participants were calibrated with a nine-point calibration. The participants

were then seated in the chair and behind the chinrest. By adjusting the height of the chair, it

was made sure that their eyebrows were below the highest bar, and that their chin rested on

the lowest bar. The participants were asked to adjust the height of the chair themselves, to make it as comfortable as possible. The keyboard had to be reachable for answering the questions and pressing the spacebar and was placed so that the participants could see the keys in the corners of their eyes without moving their head. Yellow stickers were placed on the 'z' (yes) and 'm' (no) buttons, to make sure the buttons were in sight. After that, the eye tracking experiment was run, which took about 30 minutes including the set up and calibration.

At the start of the experiment, an instruction screen appeared, in which the participants were told that they were going to read 120 sentences one at a time, and that they would have a break halfway. They were instructed to read the sentences in a natural fashion, to look at the fixation point until the next sentence appeared, to press to spacebar after reading a sentence, and to answer questions correctly when they appeared. After successful calibration, five practice sentences with two practice questions followed. After the practice questions, the participants had the opportunity to ask questions. The participant then started reading the experimental sentences. One trial consisted of a fixation point on the left side of the screen, on the exact spot where the first word of the sentence would appear. This fixation point was simultaneously a drift correction (for correcting small drifts in the gaze position calculation). For every trial, the experimenter had to accept the drift correction to go to the next trial, by pressing the spacebar when the gaze was approaching the fixation point. Then, a sentence appeared. The participants pressed the spacebar when they had read and understood the sentence. In 30 of 120 trials a comprehension question appeared, that could be answered with yes ('z') or no ('m') by pressing the keyboard.

After 60 trials, the participants were given the opportunity to have a short break. After the break, another calibration was performed, and the participants proceeded with reading the remaining 60 sentences. When the eye tracking experiment was finished, the participants filled out a language background questionnaire, which listed all the languages the participants

knew and how well they knew them in writing, speaking, listening, and reading. It also mapped how much exposure they had had to those languages and, specifically, the exposure they had had to English, both in their study program as in their leisure time. The questionnaire can be found in Appendix C. The participants also filled out an English reading task: the Vernon-Warden reading test (Hedderly, 1996). The Vernon-Warden reading test is originally designed for testing English children who may have reading problems, but it is also a good tool for estimating English reading proficiency of second language learners. It consisted of 42 'fill the gap' sentences with five multiple choice answers. Participants had exactly 10 minutes to complete the questionnaire. The ultimate score is based on the number of correct answers. Depending on the number of wrong answers, a correction for guessing was deducted from the raw score, of which a percentile was calculated, indicating the reading proficiency. The total duration of the experiment was 60 minutes.

3.6 Apparatus

The eye tracker that was used is the Eyelink 1000+ (SR Research Ltd., Mississauga, Ont., Canada), with a sampling rate of maximal 1000 Hz. The experiment was built and run with the software *Experiment Builder*, developed by SR research. The illuminator power was set to 75% percent, and there was an automatic threshold for the pupil and corneal reflection. There were two computers in the lab room: a display computer for the experiment to run on, and a host PC for the eye tracker to work on. The monitor of the display computer was in the booth, while the host PC monitor and the computers were outside the booth, on the desk of the experimenter. The display PC monitor was of the brand BENQ, type xl2420t (resolution: 1920 x 1080), the host PC monitor was from Acer, and the computers were from DELL, both running on Windows 7. To look at the data and to prepare it for further analyses *Fixation* (version 0.1.0.29; Cozijn, 2006) was used.

The experiment was conducted in one of the lab rooms at the CLS lab in the Erasmus

building of the Radboud University Nijmegen. It contained a sound-proof eye-tracking booth, a table with a chair, and a desk with the monitors and screens for the eye tracker. The booth was about 2,5m by 4m in size. Inside the booth, the eye tracker was placed on a desk, and in front of the monitor on which the experiment was run. The participants were seated on a chair that could be locked, and the participants' head was placed on a chin rest exactly in front of the tracker. The distance between the eye tracker and the chin rest was 50cm. In the booth, a strip of LED lights was placed in the corners against the ceiling, surrounding the whole booth, ensuring a calm, but bright lighting. The light could be adjusted from 1 to 12, and was always set to 8 during the experiment.

3.7 Data analysis

To analyse the data, several steps were carried out. To transfer the results from the eye-tracker to *Fixation* (Cozijn, 2006), the EDF-files retrieved from the *Experiment Builder* software were converted into ASC-files. This specific version of *Fixation* was adapted to read files from *Experiment Builder* software, from an Eyelink 1000+. The areas of interest were extracted from the folder 'runtime' in the experiment folder. This contained both the areas of interest and the images of the sentences shown during the experiment. In Fixation, these were merged together with the locations and times of every fixation per sentence per participant. Only the target sentences were analysed.

By means of the program *Fixation*, fixations were assigned to areas of interest, to determine the exact reading times per area of interest. For text, the program does this already automatically for the fixations that fall within these areas. This way of pre-processing eye tracking data ensures that when the participant has moved, or when the calibration has been lost, the fixations can be reassigned to another area of interest that fits the most probable reading behaviour of the participant. For every sentence the fixations were checked to see whether there was a certain pattern, i.e., whether the participants read fast or slow, whether he

or she read the sentence multiple times or only once, whether there were a lot of regressions or blinks and whether there were fixations too far away from the areas of interest.

When a fixation was significantly deflected from the text and areas of interest, the fixation was not assigned and excluded from the analyses. This could have been caused by movements of the participants or calibration/drift correct errors. Fixations that were not precisely fixed on a word, but were above, beneath, right, or left of it, were assigned in the way that fitted the reading direction, speed, and behaviour the best. Blinks were also taken into consideration. A blink was preceded and followed by a fixation, one before the eye closed, and one after the eye had opened again. Blinks that fell in an area of interest based on the preceding and following fixations were included in the analysis, since language processing continues during saccades and blinks (Irwin, 1998).

3.8 Statistical analysis

The statistical analysis was performed via SPSS (version 24). A mixed ANOVA was conducted with Grammaticality and Semantic Bias as within-subjects factors and Group as a between-subjects factor. Based on the outcomes of the first analysis, another mixed ANOVA was performed with Grammaticality as within-subjects factor, and Amount of Exposure and Group as between-subjects factors. In both analyses, first pass gaze duration was the dependent variable. Only the verbs of the third verb phrase and the non-verb parts of the third verb phrase were analysed. The verbs of the first verb phrase were also analysed as a control factor, as these were the same between the grammatical and ungrammatical conditions. In addition, item analyses were performed with semantic bias as a between-items factor, because of small sample size.

Before the statistical analysis could take place, some changes were made to the dataset. Firstly, the data of the participants and stimuli that were not suitable for analysis were excluded. When a participant answered more than 40% of the comprehension questions

wrong, his/her data were excluded from the statistical analyses. This did not occur. Secondly,

the reading times that were three SD's above or under the mean were excluded from analysis.

This was 8.5% for the analysis of the reading times of the verb in VP1, 7.6% for the analysis

of the reading times of the verb in VP3, and 1.4% for the non-verb parts of VP3.

4. Results

In each paragraph, the results of three important areas from the sentences will be

discussed: the verbs in VP1, the verbs in VP3 and the non-verb parts of VP3.

4.1 Semantic bias, Grammaticality and Language group

Table 3 shows the means and SE's of the reading times in milliseconds per Group,

Semantic Bias and, Grammaticality for the verbs in VP1, verbs in VP3, and the non-verb parts

in VP3.

Table 3

Mean and SE's of the reading times (ms) for the verbs in VP1, the verbs in VP3, and the non-
verb parts in VP3.

Group	Bias	Grammaticality	Verbs VP1		Verbs VP3		Non-verb parts VP3	
			M	*SE*	*M*	*SE*	*M*	*SE*
Dutch	Semantic	Grammatical	489	24	320	25	600	42
		Ungrammatical	533	25	371	24	683	45
	Neutral	Grammatical	368	20	321	25	821	59
		Ungrammatical	348	21	299	22	996	73
English	Semantic	Grammatical	500	27	338	29	621	47
		Ungrammatical	453	28	380	27	627	51
	Neutral	Grammatical	316	23	352	28	917	67
		Ungrammatical	315	24	281	24	1009	83

4.1.1. Verbs in VP1. There was a significant main effect of Semantic Bias on reading time, $F_1(1, 39) = 173.491$, $p <.001$, $\eta_p^2 = .816$, $F_2(1, 22) = 85.519$, $p <.001$, $\eta_p^2 = .795$, $\min F'(1, 43) = 57.283$, $p < .001$, suggesting that sentences without a semantic bias have faster reading times ($M = 337$, $SE = 13$) than sentences with a semantic bias ($M = 494$, $SE = 15$). There was no significant effect of Grammaticality on reading time, $F_1 < 1$, $F_2 < 1$, suggesting there was no difference between the grammatical and ungrammatical condition. There was no significant effect of Group on reading time, $F_1(1, 39) = 2.363$, $p = .132$, $\eta_p^2 = .057$, $F_2(1, 22) = 5.224$, $p = .032$, $\eta_p^2 = .192$, $\min F'(1, 60) = 1.627$, $p = .207$, suggesting there was no difference between the English and the Dutch language group. There was no significant interaction between Semantic bias and Group, $F_1 < 1$, $F_2 < 1$, indicating that the effect of semantic bias on reading times did not depend on different language group. There was no significant interaction between Grammaticality and Group, $F_1(1 ,39) = 1.406$, $p = .243$, $\eta_p^2 = .035$, $F_2(1, 22) = 1.036$, $p = .320$, $\eta_p^2 = .045$, $\min F' < 1$, suggesting that grammaticality did not depend on language group. There was no significant interaction between Semantic Bias and Grammaticality, $F_1 < 1$, $F_2 < 1$, suggesting that the effect of semantic bias did not depend on grammaticality. There was no significant interaction between Group, Semantic bias and Grammaticality, $F_1(1, 39) = 4.053$, $p = .051$, $\eta_p^2 = .094$, $F_2(1, 22) = 1.036$, $p = .320$, $\eta_p^2 = .045$, $\min F' < 1$.

4.1.2. Verbs in VP3. There was a significant main effect of Semantic Bias on reading time, $F_1(1, 39) = 5.756$, $p = .021$, $\eta_p^2 = .129$, $F_2(1, 22) = 1.708$, $p = .205$, $\eta_p^2 = .072$, $\min F'(1, 35) = 1.318$, $p = .259$, indicating that sentences without a semantic bias had lower reading times ($M = 313$, $SE = 13$) than sentences with a semantic bias ($M = 352$, $SE = 16$). There was no significant main effect of Group, $F_1 < 1$, $F_2 < 1$, indicating that there was no difference between the English and Dutch group when comparing reading time. There was no significant main effect of Grammaticality on reading time, $F_1 < 1$, $F_2 < 1$, which means there was no

difference between the two grammaticality conditions. There was a significant interaction between Semantic Bias and Grammaticality, $F_1(1, 39) = 10.133$, $p = .003$, $\eta_p^2 = .206$, $F_2(1, 22) = 6.309$, $p = .020$, $\eta_p^2 = .223$, $minF'(1, 48) = 3.889$, $p = .054$, and the means and SE's suggest that in semantically neutral sentences, the ungrammatical sentences were read faster ($M = 290$, $SE = 16$) than the grammatical sentences ($M = 336$, $SE = 18$), while in the semantically biased sentences, the ungrammatical sentences ($M = 376$, $SE = 18$) were read slower than the grammatical sentences ($M = 329$, $SE = 19$). There was no significant interaction between Group and Semantic Bias, $F_1 < 1$, $F_2 < 1$, which means that the main effect of semantic bias did not depend on language group. There was also no significant interaction between Grammaticality and Group, $F_1 < 1$, $F_2 < 1$, indicating that grammaticality did not differ depending on language group. There was no significant interaction between Semantic Bias, Grammaticality and Group, $F_1 < 1$, $F_2 < 1$.

4.1.3. Non-verb parts of VP3. There was a significant main effect of Semantic Bias on reading time, $F_1(1, 39) = 108.878$, $p < .001$, $\eta_p^2 = .736$, $F_2(1, 22) = 57.563$, $p < .001$, $\eta_p^2 = .723$, $minF'(1, 44) = 37.655$, $p < .001$, indicating that sentences without a semantic bias had higher reading times ($M = 936$, $SE = 42$) on this area than sentences with a semantic bias ($M = 633$, $SE = 30$). There was also a significant main effect of Grammaticality on reading time, $F_1(1, 39) = 12.230$, $p = .001$, $\eta_p^2 = .239$, $F_2(1, 22) = 21.355$, $p < .001$, $\eta_p^2 = .493$, $minF'(1, 61) = 7.776$, $p = .007$, and the means suggest that in ungrammatical sentences this area was read slower ($M = 829$, $SE = 38$) than in grammatical sentences ($M = 740$, $SE = 33$). There was no significant main effect of Group, $F_1 < 1$, $F_2 < 1$, which means that there were no differences between the English and the Dutch group. There was no significant interaction between Semantic Bias and Group, $F_1(1, 39) = 1.524$, $p = .224$, $\eta_p^2 = .038$, $F_2(1, 22) = 1.093$, $p = .307$, $\eta_p^2 = .047$, $minF' < 1$, indicating that the effect of semantic bias did not differ between English language group and the Dutch language group. There was no significant

interaction between Grammaticality and Group, $F_1(1, 39) = 2.473$, $p = .124$, $\eta_p^2 = .060$, $F_2 < 1$,

min$F' < 1$, which indicates that the main effect of grammaticality did not differ between the

English and the Dutch language group. There was no significant interaction between Semantic

Bias and Grammaticality, $F_1(1, 39) = 1.553$, $p = .220$, $\eta_p^2 = .038$, $F_2(1, 22) = 5.140$, $p = .034$,

$\eta_p^2 = 1.89$, min$F'(1, 57) = 1.192$, $p = .279$, indicating that the main effects of grammaticality

and bias were not dependent of each other. There was no significant interaction between

Semantic Bias, Grammaticality and Group, $F_1 < 1$, $F_2 < 1$.

4.2 Semantically neutral items: Exposure, Grammaticality and Language group

Table 4 shows the means and SE's of the reading times in milliseconds per Group,

Amount of exposure and, Grammaticality for the verbs in VP1, verbs in VP3, and the non-

verb parts in VP3.

Table 4

*Mean and SE's in of the reading time (ms) for the verbs in VP1, the verbs in VP3 and the non-
verb parts in VP3.*

Group	Amount of Exposure	Grammaticality	Verbs VP1		Verbs VP3		Non-verb parts VP3	
			M	*SE*	*M*	*SE*	*M*	*SE*
Dutch	1 year	Grammatical	309	99	189	114	1131	269
		Ungrammatical	186	101	212	105	762	369
	2 years	Grammatical	380	45	377	51	1094	121
		Ungrammatical	428	45	313	47	1195	165
	3 years	Grammatical	385	44	281	51	770	121
		Ungrammatical	372	45	273	47	925	165
	4 years	Grammatical	362	29	325	33	702	78
		Ungrammatical	317	29	311	30	962	106

Engels	1 year	Grammatical	324	33	302	38	895	90
		Ungrammatical	303	34	259	35	1010	123
	2 years	Grammatical	294	50	378	57	927	135
		Ungrammatical	326	50	296	52	959	184
	3 years	Grammatical	356	50	453	57	1015	135
		Ungrammatical	324	50	348	52	1092	184
	4 years	Grammatical	170	99	291	114	677	269
		Ungrammatical	341	101	142	105	869	369

4.2.1. Verbs in VP1. There was no significant main effect of Grammaticality on reading time, $F_1 < 1$, $F_2 < 1$, indicating that there was no difference between the grammatical and ungrammatical condition. There was no significant main effect of Amount of Exposure on reading time, $F_1 (3, 33) = 1.243$, $p = .310$, $\eta_p^2 = .102$, $F_2 (3, 76) = 1.412$, $p = .246$, $\eta_p^2 = .053$, min$F$' < 1, suggesting that there was no main influence of number of years that someone was exposed to English or Dutch in their studies. There was also no significant main effect of Group on reading time, $F_1 (1, 33) = 1.079$, $p = .306$, $\eta_p^2 = .032$, $F_2 < 1$, min$F$' < 1, suggesting that the Dutch and English language group did not differ. There was no significant interaction between Grammaticality and Group on reading time, $F_1 (1, 33) = 1.959$, $p = .171$, $\eta_p^2 = .056$, $F_2 < 1$, min$F$' < 1, suggesting that there were no differences in reading time between the two conditions, which was the same in the two language groups. There was no significant interaction between Grammaticality and Amount of Exposure on reading time, $F_1 (3, 33) = 1.304$, $p = .289$, $\eta_p^2 = .106$, $F_2 (3, 76) = 1.564$, $p = .205$, $\eta_p^2 = .058$, min$F$' < 1, indicating that grammaticality did not depend on the amount of exposure participants had had. There was no significant interaction between Group and Amount of Exposure on reading time, $F_1 < 1$, $F_2 (3, 76) = 3.297$, $p = .025$, $\eta_p^2 = .115$, min$F$' < 1, indicating that the language group did not depend on the amount of exposure participants had had. There was nog significant interaction

between Grammaticality, Group and Amount of Exposure, F_1 (3, 33) = 1.178, p = .333, η_p^2 = .097, $F_2 < 1$, min$F$' < 1.

4.2.2. Verbs in VP3. There was no significant main effect of Grammaticality on reading time, F_1(1, 33) = 2.603, p = .116, η_p^2 = .073, $F_2 < 1$, min$F$' < 1, suggesting there was no difference between the grammatical and ungrammatical condition. There was no significant main effect of Amount of Exposure on reading time, F_1(3, 33) = 2.214, p = .105, η_p^2 = .168, F_2(3, 76) = 2.003, p = .121, η_p^2 = .072, min$F$'(3, 96) = 1.052, p = .373, indicating that there was no difference between the four different exposure groups. There was no significant main effect of Group on reading time, $F_1 < 1$, F_2(1, 76) = 4.725, p = .033, η_p^2 = .058, min$F$' < 1, indicating that there was no difference between the Dutch and the English language group. There was no significant interaction between Grammaticality and Group, F_1(1, 33) = 1.336, p = .256, η_p^2 = .039, $F_2 < 1$, min$F$'< 1, suggesting grammaticality of a sentence did not depend on the language group in reading time. There was no significant interaction between Grammaticality and Amount of Exposure $F_1 < 1$, $F_2 < 1$, indicating that grammaticality did not depend on amount of exposure. There was no significant interaction between Group and Amount of Exposure on reading time, F_1(3, 33) = 2.284, p = .097, η_p^2 = .172, F_2(3, 76) = 1.103, p = .353, η_p^2 = .041, min$F$' < 1. suggesting that the reading times of language groups did not depend on amount of exposure. There was no significant interaction between Grammaticality, Group and Amount of Exposure on reading time, $F_1 < 1$, $F_2 < 1$.

4.2.3. Non-verb parts in VP3. There was no significant main effect of Grammaticality on reading time, $F_1 < 1$, F_2(1, 76) = 1.527, p = .220, η_p^2 = .019, min$F$' < 1, suggesting there was no difference between the grammatical and ungrammatical condition. There was no significant main effect of Amount of Exposure on reading time, $F_1 < 1$, $F_2 < 1$, suggesting there was no difference between the four exposure groups. There was no significant main effect of Group on reading time, $F_1 < 1$, $F_2 < 1$, suggesting there was no

difference between the English and the Dutch language group. There was no significant interaction between Grammaticality and Group on reading time, $F_1 < 1$, $F_2 < 1$, indicating that grammaticality did not depend on language group. There was no significant interaction between Grammaticality and Amount of Exposure on reading time, $F_1 < 1$, $F_2(3,\ 76) = 1.709$, $p = .172$, $\eta_p^2 = .060$, $minF' < 1$, indicating that grammaticality did not depend on the amount of exposure. There was no significant interaction between Group and Amount of Exposure on reading time, $F_1 < 1$, $F_2(3, 76) = 3.361$, $p = 0{,}23$, $\eta_p^2 = .112$, $minF' < 1$, suggesting that the language group did not depend on the amount of exposure. There was nog significant interaction of Grammaticality, Group and Amount of Exposure on reading time, $F_1 < 1$, $F_2 < 1$.

5. Discussion

5.1 Summary of results

This study investigated whether the grammaticality illusion is affected by the amount of exposure to English. Two participant groups were compared: one group of Dutch native speakers with English as a second language, which were mostly exposed to Dutch in both their studies as in their free time, was compared to a group of Dutch native speakers with English and a second language, which were mostly exposed to English in both those domains. The participants also participated in an English reading test.

A first finding was an effect of semantic bias on reading times. The third verb was read faster when there was no semantic bias compared to when there was a semantic bias. This is however not surprising, as the semantically biased stimuli are completely different from the semantically neutral stimuli. This finding will therefore not be further discussed.

An important finding was a significant interaction between semantic bias and grammaticality on the reading times of the verbs in VP3. The first hypothesis can therefore be partly confirmed. Initially, it was expected that there would be an overall grammaticality

illusion. By looking at the means, it can be concluded that only in the semantically neutral condition, a grammaticality illusion occurred. In that condition, the third verb was read faster in the missing-VP2 sentences than in the grammatical ones. However, in the semantically biased stimuli, the third verb was read slower in the missing-VP2 sentences than the grammatical condition, which indicates the absence of a grammaticality illusion. Previous research (Frank et al., 2016; Vasishth et al. 2010) also concluded that when the verbs in the third verb phrase are read faster when the second verb phrase is missing from that sentence (the missing-VP2 condition), there is a grammaticality illusion. The faster reading times indicate no problems with processing and comprehending the sentence, while slower reading times do indicate problems, and probably some awareness of ungrammaticality. When looking at the reading times of the other parts of the third verb phrase, next to the verbs, there were main effects of semantic bias and grammaticality. The means show that the reading times are slower for the ungrammatical condition, than for the semantically neutral condition. The second hypothesis could not be confirmed, because there was no difference in the reading times between the Dutch language group and the English language group. Initially it was expected that in the English language group, the reading times of the third verb phrase in the missing-VP2 condition would be shorter than in the Dutch language group.

To sum up the first analysis, it seems that a grammaticality illusion occurs only in sentences without a semantic bias (semantically neutral). The semantically neutral sentences are also causing longer reading times on the last part of the sentence, the non-verb parts of the third verb phrase. The semantically biased stimuli used in the current study are based on the stimuli of Gibson and Thomas (1999), who did find a grammaticality illusion, although by measuring grammaticality judgements. Christiansen and Macdonald (2009), tested both sentences with and without a sematic bias in a self-paced reading task and found a grammaticality illusion in both sentence types. The current results do not match these

previous findings, and it is unclear how this difference between semantically biased and neutral sentences could occur.

The second analysis only tested the semantically neutral sentences, because only there a grammaticality illusion occurred. Additionally, exposure was also added as an independent variable. There were no significant differences between all conditions in both regions, i.e. the verbs of the third verb phrase, and the non-verbs of the third verb phrase. However, the means do show a certain trend that fits the third hypothesis. It was expected that in the English group it would be the case that the higher the number of years of exposure, the greater the grammaticality illusion would be. In the current results, the difference between the third verb reading times in the grammatical and ungrammatical condition increases when the amount of exposure increases. In other words, the reading times of the VP3 in the missing-VP2 condition get slower as the years of exposure to English increase. This would indicate that there is a trend that the more English exposure someone has, the more likely it is that person will experience a grammaticality illusion. Due to low statistical power, these differences did not reach significance. The sample size of the different years of exposure groups were not evenly divided. Therefore, the answer to the research question is that in this study only a trend was seen in the data that does point to an influence of amount of English exposure on the size of the effect.

5.2 Theoretical implications

Comparing the interference account, the statistical account and the working memory account with the results of the current study, no conclusive answer can be given to the question of which account fits the best. The interference account (Haüssler & Bader, 2015) explains the missing-VP2 effect by stating that there is competition for the attachment of the verb to the first of the second noun phrase. Apparently, in this study, the semantic bias in a sentence causes no interference when there are three consecutive verb phrases. However,

when there is no semantic bias, there is interference, and consequently a missing-VP2 effect. The working memory account could explain the results when assuming that it should be harder to process sentences without a semantic bias, than sentences with a semantic bias. The semantic relations in the sentence could help the participants 'remember' that there is another noun phrase that needs to be assigned a verb phrase, which prevents a grammaticality illusion. Both the statistical account (Frank et al. 2016), and the hybrid statistical and working memory account (Frank & Ernst, submitted) would state that more exposure would lead to bigger missing-VP2 effect. The current results cannot confirm these accounts, but the trend that is seen does point towards that direction. That would mean that working memory is not involved in the missing-VP2 effect, and that it would be explained by sensitivity to language statistics. In sum, there is no conclusive answer, but further research could be able to find evidence both statistical accounts.

5.3 Limitations and suggestions for further research

There are several methodological issues and limitations that need to be taken into consideration when interpreting the results of this study. Firstly, there is a lot of variation in the reading strategies that the participants have. One would read a sentence in one go, and then see whether it is necessary to look back and read (parts of) the sentence again, while another would reread a difficult section immediately, and only move on when the previous part was comprehended. Frazier and Rayner (1982) found that when participants read garden-path sentences, there are two strategies of how one could read them: through *minimal attachment* and *late closure*, and arguably, this distinction could be applied to the sentences used in the present study as well. The use of eye tracking as a method allowed the participants to look back and reread, which in turn leads to more variation in reading strategies. This could easily be controlled for by using self-paced reading as a method. However, this study deliberately chose for eye tracking, since this method allows participants to read in a more

natural way compared to self-paced reading.

Another limitation is the definition of exposure. It is a challenge to control for the participants' background in term of amount of exposure. In this study, participants were grouped together based on their study background. This is however unfortunately not a perfect separation. First of all, the fact that someone follows an English bachelor or master does not automatically imply a high proficiency or high exposure. Master programs, for example, only have a limited amount of obligatory lectures compared to regular bachelor programs. For proficiency, we also conducted a English reading test, but only to be able to exclude people with low proficiency scores. Another aspect is that in a English bachelor or master that is given in the Netherlands, it is often the case that the teachers and students that are the source of exposure are non-native speakers of English. Unsworth (2016) stated that the non-nativeness and proficiency of the source of exposure (i.e. the quality of the input) does have a major influence on the ultimate outcomes proficiency measures when testing young children. It could also be the case that some of the participants do have a language aptitude and have always been good at second languages, even if they have not done anything with for example English or French since they left high school.

A suggestion for further research is to look more closely at the influence of amount of exposure (in years) and study years. Unfortunately, this study was not able to find a significant influence due to a small sample size. The sample was not evenly divided over the four exposure groups per language group. However, when it would be possible to look at the possible influence of the amount of exposure to English with a bigger and more evenly divided sample size, it is interesting to see whether the missing-VP2 effect gets larger as the amount of exposure gets higher. The results in this study only indicated a trend in that direction. It would even be better to conduct the study longitudinally, to control the amount of exposure and to be able to see within one participant whether effect gets larger over the years,

as the amount and length of exposure to English gets higher.

To really get a good idea of the unconscious reading processes of readers of these double centre-embedded relative clauses, it would be interesting to conduct this study with EEG as a method, looking at event-related potentials (ERP's). By comparing the ERP's of the grammatical and ungrammatical condition, the actual difference of processing can be measured. If there is a real grammaticality illusion, the magnitude of the peak when reading the ungrammatical sentence would then be smaller than or equal to reading the grammatical sentence. This method could also elucidate whether the process is different or the same in an L2. Van Hell and Tokowicz (2010) reviewed L2 sentence processing studies that used ERP's, and they conclude that ERP's can help to distinguish subtle differences, that behavioural studies cannot detect.

6. Conclusion

The present study examined the influence of amount of English exposure on the grammaticality illusion. In contrast to our expectations, an interaction of semantic bias and grammaticality was found. Only in the semantically neutral sentences, a grammaticality illusion occurred. That is, the sentences missing the second verb phrase, were read faster than the grammatical condition. This would indicate that the grammaticality illusion is influenced by semantics. This is a topic that would be interesting for further research. When looking at the influence of exposure on the grammaticality illusion, only a non-significant trend was found, which indicates that the more exposure someone has, the bigger the difference is between the reading times of the grammatical and the ungrammatical condition. This was in line with the expectations, but further research should focus on supporting this notion.

Acknowledgements

I would like to thank my supervisors dr. S.L. Frank and dr. R. Cozijn for enabling me to carry out this study and helping me with learning everything. I am very grateful that they invested a lot of time and effort in making this Master's Thesis possible. I would also like to thank the CLS lab and their managers for helping during the testing period.

References

Bock, K., & Miller, C. A. (1991). Broken Agreement. *Cognitive Psychology*, *23*, 45-93**.**

Cenoz, J. (2001). The effect of linguistic distance, L2 status and age on cross-linguistic influence in third language acquisition. *Cross-linguistic influence in third language acquisition: Psycholinguistic Perspectives*, *111*, 8-20.

Christiansen, M. H., & MacDonald, M. C. (2009). A usage-based approach to recursion in sentence processing. *Language Processing, 59*, 126-161.

Cozijn, R. (2006). Het gebruik van oogbewegingen in leesonderzoek. *Tijdschrift voor Taalbeheersing, 28*, 1270-1276.

Eyelink 1000+ [Apparatus and software]. (2004). Mississauga, Ontario, Canada: SR Research.

Frank, S. L., & Ernst, P. (2017) *Judgements about double-embedded relative clauses differ between languages.* Manuscript submitted for publication.

Frank, S. L., Trompenaars, T., & Vasishth, S. (2016). Cross-linguistic differences in processing double-embedded relative clauses: Working-memory constraints or language statistics. *Cognitive Science, 40*, 1-25.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*, 178-210.

Gibson, E. (1991). *A computational theory of human linguistic processing: Memory limitations and processing breakdown.* Ph.D. thesis, Carnegie Mellon University, Pittsburgh. PA.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68,* 1-76.

Gibson, E. (2000). Dependency locality theory: A distance-based theory of linguistic complexity. In: A. Marantz, Y Miyashita, & W. O'Neil (Eds.), *Image, language, brain: Papers from the first mind articulation project symposium* (pp. 95-126). Cambridge, MA: MIT Press.

Gibson, E., & Thomas, J. (1999). Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes, 14*, 225-248.

Gimenes, M., Rigalleau, F., & Gaonac'h, D. (2009). When a missing verb makes a French sentence more acceptable. *Language and Cognitive Processes, 24*, 440-449.

Häussler, J., & Bader, M. (2015). An interference account of the missing-VP effect. *Frontiers in Psychology*, *6*.

Hedderly, R. (1996). Vernon-Warden Reading Test, Restandardised 1993 and 1994. *Dyslexia Review*, *7*, 11-16.

Irwin, D. E. (1998). Lexical processing during saccadic eye movements. *Cognitive Psychology, 36*, 1-27.

Koster, J. (2016). *A matter of omission: A study on the comprehension and acceptability of doubly centre-embedded sentences in English* (Bachelor's thesis, Radboud University Nijmegen).

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods, 44*, 325-343.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental*

    Psychology, *64*, 482-488.

Ojima, S., Matsuba-Kurita, H., Nakamura, N., Hoshino, T., & Hagiwara, H. (2011). Age and

    amount of exposure to a foreign language during childhood: Behavioral and ERP data

    on the semantic comprehension of spoken English by Japanese children. *Neuroscience*

    Research, *70*, 197-205.

Polinsky, M. (2012). Headedness, again. *UCLA Working Papers in Linguistics, 17*, 348-359.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of

    research. *Psychological Bulletin*, *124*, 372-422.

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of

    comprehension processes in reading. *Scientific Studies of Reading*, *10*, 241-255.

Saville-Troike, M. (2013). *Introducing second language acquisition* (2nd ed.). Cambridge,

    United Kingdom: Cambridge University Press.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention,*

    Perception, & Psychophysics, *74*, 5-35.

Unsworth, S. (2016). Quantity and quality of language input in bilingual language

    development. In E. Nicoladis & S. Montanari (Eds.), *Lifespan perspectives on*

    bilingualism (pp. 136-196). Mouton de Gruyter/APA.

Van Hell, J. G., & Tokowicz, N. (2010). Event-related brain potentials and second language

    learning: Syntactic processing in late L2 learners at different L2 proficiency

    levels. *Second Language Research*, *26*, 43-74.

Vasishth, S., Suckow, K., Lewis, R. L., & Kern, S. (2010). Short-term forgetting in sentence

    comprehension: Crosslinguistic evidence from verb-final structures. *Language and*

    Cognitive Processes, 25, 533-567

Vigliocco, G., & Nicol, J. (1998). Separating hierarchical relations and word order in

language production: Is proximity concord syntactic or linear? *Cognition*, *68*, 13-29.

Appendices

Appendix A

Stimuli English 2 conditions

This appendix shows the stimuli in both conditions, with the three different verb phrases, and the VP2 condition in which the second verb phrase is omitted. It also shows where the sentence was divided and which part was on the second text line, indicated by [/].

Semantically biased

1) The book that the student who the new catalog [had confused a great deal]V1 [was studying / in the library]V2 [was missing an important page.]V3
   a. The book that the student who the new catalog [had confused / a great deal]V1 [was missing an important page.]V3
2) The lullaby that the singer who the record label [had signed to a big contract]V1 [was singing / yesterday]V2 [was written sixty years ago.]V3
   a. The lullaby that the singer who the record label [had signed to a big / contract]V1 [was written sixty years ago.]V3
3) The game that the child who the lawnmower [had startled in the yard]V1 [was playing in / the morning]V2 [lasted for several hours.]V3
   a. The game that the child who the lawnmower [had startled in / the yard]V1 [lasted for several hours.]V3
4) The crime that the gangster who the story [had profiled thoroughly]V1 [had planned / for weeks]V2 [was solved in the middle of the night.]V3
   a. The crime that the gangster who the story [had profiled / thoroughly]V1 [was solved in the middle of the night.]V3
5) The picture that the student who the school [had expelled for cheating]V1 [was hurriedly / copying]V2 [was printed in a popular magazine.]V3
   a. The picture that the student who the school [had expelled for / cheating]V1 [was printed in a popular magazine.]V3
6) The trophy that the athlete who the restaurant [had hired as a spokesman]V1 [had won at the / championship]V2 [was stolen the day after.]V3
   a. The trophy that the athlete who the restaurant [had hired as a / spokesman]V1 [was stolen the day after.]V3
7) The apartment that the maid who the service [had sent over]V1 [was cleaning every / week]V2 [was decorated with beautiful flowers.]V3
   a. The apartment that the maid who the service [had sent / over]V1 [was decorated with beautiful flowers.]V3
8) The shirt that the seamstress who the officer [had investigated last week]V1 [was carefully / mending]V2 [needed to be washed at thirty degrees.]V3
   a. The shirt that the seamstress who the officer [had investigated / last week]V1 [needed to be washed at thirty degrees.]V3
9) The lecture that the professor who the newspaper [had interviewed in detail]V1 [was teaching / poorly]V2 [was attended by twenty students.]V3

a. The lecture that the professor who the newspaper [had interviewed in / detail]V1 [was attended by twenty students.]V3

10) The novel that the horror author who the publisher [had fired recently]V1 [had typed / quickly]V2 [was banned by the local library.]V3

   a. The novel that the horror author who the publisher [had fired / recently]V1 [was banned by the local library.]V3

11) The prayer that the monk who the religious man [had persecuted fiercly]V1 [was chanting / every day]V2 [was echoing in the empty church.]V3

   a. The prayer that the monk who the religious man [had persecuted / fiercly]V1 [was echoing in the empty church.]V3

12) The play that the actor who the company [had underpaid repeatedly]V1 [was performing last / month]V2 [was extremely well written by a famous author.]V3

   a. The play that the actor who the company [had underpaid / repeatedly]V1 [was extremely well written by a famous author.]V3

Semantically neutral

1) The carpenter who the craftsman who the peasant [carried a long way]V1 [hurt on / purpose]V2 [supervised the apprentice in the garden.]V3

   a. The carpenter who the craftsman who the peasant [carried a long / way]V1 [supervised the apprentice in the garden.]V3

2) The mother who the daughter who the sister [found within minutes]V1 [frightened with / a mask]V2 [greeted the grandmother on the tricycle.]V3

   a. The mother who the daughter who the sister [found within / minutes]V1 [greeted the grandmother on the tricycle.]V3

3) The worker who the tenant who the foreman [looked for today]V1 [injured / with a knife]V2 [questioned the shepherd in the office.]V3

   a. The worker who the tenant who the foreman [looked / for today]V1 [questioned the shepherd in the office.]V3

4) The trader who the businessman who the professor [hired for a certain period]V1 [confused / with questions]V2 [annoyed the investor in the morning.]V3

   a. The trader who the businessman who the professor [hired for a certain / period]V1 [annoyed the investor in the morning.]V3

5) The painter who the musician who the father [missed by ten minutes]V1 [sheltered in / the attic]V2 [cooked for the artist in the kitchen.]V3

   a. The painter who the musician who the father [missed by ten / minutes]V1 [cooked for the artist in the kitchen.]V3

6) The saxophonist who the trumpeter who the conductor [brought along]V1 [distracted / with flowers]V2 [thanked the violinist in his speech.]V3

   a. The saxophonist who the trumpeter who the conductor [brought / along]V1 [thanked the violinist in his speech.]V3

7) The pharmacist who the optician who the stranger [saw from a distance]V1 [troubled with / some remarks]V2 [questioned the customer at the counter.]V3

   a. The pharmacist who the optician who the stranger [saw from / a distance]V1 [questioned the customer at the counter.]V3

8) The cleaner who the janitor who the doctor [recognized by his uniform]V1 [sought / after lunch]V2 [surprised the patient in the hallway.]V3

    a. The cleaner who the janitor who the doctor [recognized by his / uniform]V1 [surprised the patient in the hallway.]V3

9) The dancer who the singer who the bystander [admired with jealousy]V1 [met at / the party]V2 [tipped the doorman at the door.]V3

    a. The dancer who the singer who the bystander [admired with / jealousy]V1 [tipped the doorman at the door.]V3

10) The artist who the sportsman who the guard [shouted at yesterday]V1 [annoyed / with boring stories]V2 [instructed the newscaster in the studio.]V3

    a. The artist who the sportsman who the guard [shouted / at yesterday]V1 [instructed the newscaster in the studio.]V3

11) The son who the father who the teacher [saw for the first time]V1 [disturbed by / barging in]V2 [visited the grandfather in the nursing home.]V3

    a. The son who the father who the teacher [saw for the first / time]V1 [visited the grandfather in the nursing home.]V3

12) The defence who the prosecutor who the spy [looked at for several minutes]V1 [surprised / by crying]V2 [convinced the judge in the courtroom.]V3

    a. The defence who the prosecutor who the spy [looked at for several / minutes]V1 [convinced the judge in the courtroom.]V3

Appendix B

Changes in Stimuli

This appendix shows the changes that were made to the original stimuli, to make them fit better into the experiment.

Semantically biased

1.	The book that the student who the new catalog [had confused a great deal]V1 [was studying in the library]V2 [was missing an important page.]V3	The ~~ancient manuscript~~ book that the ~~grad~~ student who the new ~~card~~ catalog had confused a great deal was studying in the library was missing a<u>n important</u> page.
2.	The lullaby that the singer who the record label [had signed to a big contract]V1 [was singing yesterday]V2 [was written sixty years ago.]V3	The lullaby that the ~~famous country~~ singer who the record label had signed to a big contract was singing yesterday was written ~~seventy~~ <u>sixty</u> years ago.
3.	The game that the child who the lawnmower [had startled in the yard]V1 [was playing in the morning]V2 [lasted for several hours.]V3	The game that the child who the lawnmower had startled in the yard was playing in the morning lasted for <u>several</u> hours.
4.	The crime that the gangster who the story [had profiled thoroughly]V1 [had planned for weeks]V2 [was solved in the middle of the night.]V3	The crime that the gangster who the story had profiled <u>thoroughly</u> had planned for weeks was ~~quickly~~ solved <u>in the middle of the night.</u>
5.	The picture that the student who the school [had expelled for cheating]V1 [was hurriedly copying]V2 [was printed in a popular magazine.]V3	The picture that the ~~artist~~ <u>student</u> who the school had expelled for cheating was hurriedly copying was printed in a <u>popular</u> magazine.
6.	The trophy that the athlete who the restaurant [had hired as a spokesman]V1 [had won at the championship]V2 [was stolen the day after.]V3	The trophy that the athlete who the restaurant had hired as a spokesman had won at the ~~track~~ <u>championship</u> was stolen ~~later~~ <u>the day after</u>.
7.	The apartment that the maid who the service [had sent over]V1 [was cleaning every week]V2 [was decorated with beautiful flowers.]V3	The apartment that the maid who the service had sent over was cleaning every week was ~~well~~ decorated <u>with beautiful flowers</u>.
8.	The shirt that the seamstress who the officer [had investigated last week]V1 [was carefully mending]V2 [needed to be washed at thirty degrees.]V3	The shirt that the seamstress who the ~~immigration~~ officer had investigated last week was carefully mending needed to be ~~dry cleaned~~ <u>washed at thirty degrees</u>.
9.	The lecture that the professor who the newspaper [had interviewed in	The lecture that the professor who the newspaper ~~story~~ had ~~just profiled~~

	detail]V1 [was teaching poorly]V2 [was attended by twenty students.]V3	interviewed in detail was teaching poorly was ~~not well~~ attended by twenty students.
10.	The novel that the horror author who the publisher [had fired recently]V1 [had typed quickly]V2 [was banned by the local library.]V3	The novel that the horror author who the ~~publishing company~~ publisher had ~~fired~~ recently ~~hired~~ had typed quickly was banned by the local library.
11.	The prayer that the monk who the religious man [had persecuted fiercly]V1 [was chanting every day]V2 [was echoing in the empty church.]V3	The prayer that the monk who the religious ~~fanatic~~ man had persecuted ~~relentlessly~~ fiercly was chanting every day was echoing in the empty church.
12.	The play that the actor who the company [had underpaid repeatedly]V1 [was performing last month]V2 [was extremely well written by a famous author.]V3	The ~~monologue~~ play that the actor who the ~~movie industry~~ company had ~~snubbed~~ underpaid repeatedly was performing last month was extremely well written by a famous author.

Semantically neutral

13.	The carpenter who the craftsman who the peasant [carried a long way]V1 [hurt on purpose]V2 [supervised the apprentice in the garden.]V3	The carpenter who the craftsman who the peasant carried a long way hurt on purpose supervised the apprentice in the garden.
14.	The mother who the daughter who the sister [found within minutes]V1 [frightened with a mask]V2 [greeted the grandmother on the tricycle.]V3	The mother who the daughter who the sister found within minutes frightened with a mask greeted the grandmother on the tricycle.
15.	The worker who the tenant who the foreman [looked for today]V1 [injured with a knife]V2 [questioned the shepherd in the office.]V3	The worker who the tenant who the foreman looked for today injured with a knife questioned the shepherd in the office.
16.	The trader who the businessman who the professor [hired for a certain period]V1 [confused with questions]V2 [annoyed the investor in the morning.]V3	The trader who the businessman who the professor hired for a certain period confused with questions annoyed the investor in the morning.
17.	The painter who the musician who the father [missed by ten minutes]V1 [sheltered in the attic]V2 [cooked for the artist in the kitchen.]V3	The painter who the musician who the father missed by ten minutes sheltered in the attic cooked for the artist in the kitchen.
18.	The saxophonist who the trumpeter who the conductor [brought along]V1 [distracted with flowers]V2 [thanked the violinist in his speech.]V3	The saxophonist who the trumpeter who the conductor brought along distracted with flowers thanked the violinist in his speech.

19.	The pharmacist who the optician who the stranger [saw from a distance]V1 [troubled with some remarks]V2 [questioned the customer at the counter.]V3	The pharmacist who the optician who the stranger saw <u>from a distance</u> troubled <u>with some remarks</u> questioned the customer at the counter.
20.	The cleaner who the janitor who the doctor [recognized by his uniform]V1 [sought after lunch]V2 [surprised the patient in the hallway.]V3	The cleaner who the janitor who the doctor recognized <u>by his uniform</u> ~~hurt~~ <u>sought after lunch</u> surprised the patient in the hallway.
21.	The dancer who the singer who the bystander [admired with jealousy]V1 [met at the party]V2 [tipped the doorman at the door.]V3	The dancer who the singer who the bystander admired <u>with jealousy</u> ~~hurt~~ <u>met at the party</u> tipped the doorman at the door.
22.	The artist who the sportsman who the guard [shouted at yesterday]V1 [annoyed with boring stories]V2 [instructed the newscaster in the studio.]V3	The artist who the sportsman who the guard shouted at <u>yesterday</u> annoyed <u>with boring stories</u> instructed the newscaster in the studio
23.	The son who the father who the teacher [saw for the first time]V1 [disturbed by barging in]V2 [visited the grandfather in the nursing home.]V3	The son who the father who the teacher saw <u>for the first time</u> disturbed <u>by barging in</u> visited the grandfather in the nursing home.
24.	The defence who the prosecutor who the spy [looked at for several minutes]V1 [surprised by crying]V2 [convinced the judge in the courtroom.]V3	The defence who the prosecutor who the spy looked at <u>for several minutes</u> surprised <u>by crying</u> convinced the judge in the courtroom.

Deleted sentences from Frank et al. (2016)

11. The clerk who the bureaucrat who the visitor forgot about [helped] annoyed the neighbor at the town hall.

13. The conductor who the choirmaster who the worker ignored [hit] berated the musician at the festival.

15. The cousin who the brother who the peasant described [pleased] hated the uncle from the farm.

16. The painter who the musician who the friend liked [disturbed] admired the poet in the pyjamas.

Appendix C

Language Background Questionnaire

Proefpersoonnummer _____

Vragenlijst talenkennis en studieachtergrond

Geboortedatum (DD/MM/JJ): _____ Geslacht: M / V

Dominante oog: _____ Bril/Lenzen?_____

Studie (Bachelor): _____

Begin- en eindjaar_____

Studie (Master indien van toepassing)_____

Beginjaar- en eindjaar _____

Huidig studiejaar_____

In welke regio(s) en land(en) heb je tot je zesde gewoond? _____

Welke talen (inclusief dialecten) spraken je ouders of verzorgers tegen je? _____

Geef aan welke talen (inclusief dialecten) je kent en hoe goed je jezelf acht in het spreken,
luisteren, lezen en schrijven in elke taal; op een schaal van 1 tot 7 (1 = zeer slecht; 4 =
middelmatig; 7 = vloeiend). Geef hier ook aan welke taal of talen je beschouwt als je
moedertaal (de taal of talen die je als eerste hebt geleerd).

Taal	Spreken	Luisteren	Lezen	Schrijven	Moedertaal?
Nederlands					
Engels					

Geef aan hoe vaak je een taal gebruikt om te spreken, luisteren, lezen en schrijven; op een
schaal van 1 tot 7 (1 = vrijwel nooit, 4 = de helft van de tijd, 7 = vrijwel altijd)

Taal	Spreken	Luisteren	Lezen	Schrijven
Nederlands				
Engels				

<u>Bij onderstaande vragen is het de bedoeling een zo goed mogelijke schatting te geven.</u>

Op welke leeftijd ben je voor het eerst in aanraking gekomen met het Engels, bijvoorbeeld via familie of vrienden, op vakantie, of in de media? ____

Op welke leeftijd ben je begonnen Engels te leren op school? ____

Studie (huidige of zojuist afgeronde)

Hoeveel uur in de week heb je college in het Engels? _____ uur

Hoeveel uur in de week heb je college in het Nederlands? _____uur
Hoeveel uur in de week lees je Engels voor je studie? _____uur

Hoeveel uur in de week lees je Nederlands voor je studie? _____uur

Vul in wat van toepassing is:

Hoeveel vakken heb je in het Engels in je huidige studiejaar? _____ vakken
Vrije Tijd

Hoeveel uur in de week luister je of praat je Engels? _____ uur

Hoeveel uur in de week lees je in je vrije tijd Engels?_____ uur

Hoeveel uur in de week lees je in je vrije tijd Nederlands? _____uur

Lezen

Hoeveel procent van wat je leest voor je studie, lees je in het Engels? _____ procent

Hoeveel procent van wat je leest voor je studie, lees je in het Nederlands? _____ procent

Hoeveel procent van wat je leest in je vrije tijd, lees je in het Engels? _____ procent
Hoeveel procent van wat je leest in je vrije tijd, lees je in het Nederlands? _____ procent