Anna Irene Andresen

Filler-Gap Processing in L2 English: Active and Island-sensitive

A study in second language processing

Master's thesis in English Linguistics
Supervisor: Dave Kush

June 2020

Master's thesis

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

NTNU
Norwegian University of
Science and Technology

Anna Irene Andresen

Filler-Gap Processing in L2 English: Active and Island-sensitive

A study in second language processing

NTNU

Norwegian University of
Science and Technology

Abstract

First Language (L1) and Second Language (L2) processing may differ from each other in many ways. It has been argued that second language users can only construct 'shallow' structural representations during incremental language processing (Clahsen & Felser, 2006). Other studies suggest that properties of one's native language may (inappropriately) transfer to L2 processing (e.g. Kim, Baek & Tremblay, 2015). The current study explores these proposals by investigating how Norwegian L1-English L2 speakers process filler-gap dependencies and whether they respect island constraints in L2 English. Results of an offline acceptability judgment study and an online self-paced reading experiment show that when reading English, L1 Norwegian participants (a) actively fill gaps in grammatical locations and (b) do not try to fill gaps inside an island domain. These findings indicate that L2 speakers are able to build rich structural representations during online sentence processing, similar to L1 speakers and that some of Norwegians' L1 'island-insensitivity' does not transfer to L2 English.

Sammendrag

Språkprosessering i ens førstespråk (L1) og andrespråk (L2) kan variere på flere måter. Det har tidligere blitt argumentert for at andrespråksbrukere kun kan konstruere overfladiske (*shallow*) strukturelle representasjoner under (stegvis) språkprosessering (Clahsen & Felser, 2006). Andre funn viser til at enkelte egenskaper ved ens førstespråk kan overføres (*transfer*) negativt til L2-prosessering. Disse teoriene blir utforsket ved å undersøke hvordan norske morsmålstalere prosesserer engelske *filler-gap dependencies*. Her ser vi på om de respekterer såkalte *island constraints* i L2-engelsk, da norsk er et språk med få island constraints sammenlignet med engelsk. Gjennom resultatene vi har hentet fra en Acceptability Judgment Task og et Self-Paced Reading-eksperiment tyder det på at morsmålstalere av norsk (a) aktivt benytter seg av en *Active-Gap Filling Strategy* ved L2-prosessering og (b) unngår å fylle in *gaps* dersom disse befinner seg i en *island domain*. Funnene fra denne studien indikerer at andrespråksbrukere kan bygge strukturelle representasjoner på lik linje med førstespråksbrukere, og at norske L1-brukere unngår å overføre egenskaper fra islands i norsk til engelsk ved prosessering av engelske filler-gap dependencies.

Acknowledgements

First, I would like to thank the fifty-seven people who took their time to participate in this experiment. Many of you chose to share this experiment with others as well. Needless to say, this project could not have been completed without you.

I would also like to express my gratitude to my supervisor, Dave Kush, for always giving me valuable feedback that I could learn from, and encouraging me whenever I felt that I could not finish this project. I have learned so much about L2 processing, islands, and experimental design this year, and working on this thesis with his guidance has really helped me grow. He also helped me conduct the statistical analysis, which I am especially grateful for.

I would also like to thank my friends and family for helping me with various tasks when writing this thesis. In particular, I would like to thank Isak for helping me create experimental items for the Self-Paced Reading task. I also owe my sister, Desiré, a huge thanks for helping me proofread this thesis.

Lastly, I would like to thank Halvard, for always encouraging and believing in me whenever I felt discouraged. Given the unusual circumstances this semester, I never thought that I would be able to finish this thesis in time for graduation. With your support, we made it happen!

Table of Contents

Figures and Tables

Abbreviations

L1 Native/First Language

L2 Second Language

SSH Shallow Structure Hypothesis

RC Relative Clause

AGF Active Gap-Filling

CP Complement Phrase

N Noun (head of Noun Phrase)

DP Determiner Phrase

TP Tense Phrase

PG Parasitic Gap

SPR Self-Paced Reading

1 Introduction

1.1 Background and motivation

A central question in Second Language (L2) research explores to what extent L2 users have the same abstract representations as native (L1) speakers and whether processes for computing those representations are similar. In this thesis, we investigate two ways in which L2 processing has been proposed to differ from L1 processing: First, in terms of representational depth and second, in terms of whether L1 processes/knowledge influence the behavior in L2. We explore these two questions by looking at the online processing of islands during filler-gap resolution.

Clahsen & Felser (2006) proposed the *Shallow Structure Hypothesis* (SSH), stating that L2 users can only create shallow representations that lack detailed syntactic information during real-time language processing. In essence, the SSH indicates that L2 speakers lack the ability to make use of grammatical information during (online) sentence processing. Recognizing syntactic islands, or phrases that do not allow extraction of syntactic constituents (Ross, 1967), requires the parser to build rich structural representations during sentence processing. As islands are complex syntactic constituents, the SSH predicts that L2 speakers should not display sensitivity to them. This thesis investigates this prediction of the SSH by testing L2 users' sensitivity to island constraints during filler-gap dependency formation. Our results show that proficient L2 speakers can follow native-like parsing strategies when processing filler-gap dependencies and that they show sensitivity to island constraints in their L2. These results are consistent with previous findings (e.g. Omaki & Schulz, 2011; Witzel, Witzel & Nicol, 2012). Overall, these findings support the view that L2 speakers are, in fact, able to build abstract structural representations during real-time processing.

It is also reported effects of native language influence, or transfer of L1 elements or patterns to the target language, in advanced L2 processing (e.g. Roberts, Gullberg & Indefrey, 2008; Kim et al., 2015). As transfer of features from one's native language has been observed in L2 processing, it is possible that transfer of L1 syntactic features happens in L2 processing of filler-gap dependencies with island domains. Particularly if the L1 does not have a specific island effect, but the L2 does.

We test the possibility of L1 transfer. Norwegian and the other mainland Scandinavian languages are known to allow movement into certain islands which are considered unacceptable in English. Research using offline judgments on island constraints in the Scandinavian languages has been conducted (e.g. Engdahl, 1983; Kush, Lohndal & Sprouse, 2018; Christensen & Nyvad, 2014). However, not much work has been done on real-time integration of English island constraints by Norwegian native speakers. The present study tests whether native Norwegian speakers attempt to fill gaps inside a particular type of island in their L2 English: a RC situated inside a subject phrase.

1.2 Research questions

The current study investigates how advanced L2 speakers generate structural representations during processing of filler-gap dependencies, and whether the L2 users are sensitive to detailed syntactic information, namely island constraints. Investigating the nature of linguistic representations generated by L2 users will give us information of possible differences between native and non-native language processing. In this thesis, three possible outcomes will be taken into consideration:

(i) L1 and advanced L2 processing are qualitatively similar.
(ii) Advanced L2 processing differs from L1 processing in that L2 speakers are not able to build rich structural representations during real-time processing.
(iii) Advanced L2 processing differs from L1 processing in that the grammatical properties of the L1 can transfer during real-time processing.

1.3 Overview of the thesis

In chapter 2, we will first present theoretical background relevant for this thesis. This includes relevant information about island constraints, active gap filling and theories of L2 processing. We will then provide a description of the method used in chapter 3. This chapter includes theoretical justification of the experimental design used to investigate the research question posed. Additionally, a description and explanation of the experiment are provided including materials, procedure and predictions of the experiment. Chapter 4 introduces the statistical analyses used to interpret the data for this study, followed by a description of the results. The implications on L2 processing suggested by our results are then discussed in chapter 5. A discussion of the reliability of the study will also be presented, referring to issues that could interfere with the results. The thesis is concluded in chapter 6.

2 Theoretical Background

2.1 Grammar of Filler-Gap Dependencies

2.1.1 Basic Filler-Gap Dependencies

This thesis explores filler-gap dependencies and how second language users process them. Filler-gap dependencies are constructions where an argument or adjunct is far from the position where it is interpreted (Clifton & Frazier, 1989, p. 274). Take for instance the sentence in (1):

(1) Which book did Mary like?

It is clear on an intuitive level that the wh-phrase *which book* is connected with the transitive verb *like* because it is interpreted as the direct object of the verb. Under a transformational analysis, the wh-phrase is analyzed as being originally generated in object position after *like*, and then moved to the front of the sentence to satisfy language specific transformational rules (Chomsky, 1957). To illustrate where the displaced phrases originated, the sentences are often written as in (2), where the underscore indicates the displaced phrase's base position:

(2) *Which book* did Mary like __?

The displaced phrases are referred to as fillers (marked in italics), and the position where the filler originated is called the gap (Fodor, 1978).

Filler-gap dependencies are harder to parse than sentences written in base form (e.g. Mary liked the book). This is because the filler has to be temporarily stored in working memory until the gap has been located (Gibson, 1998, p. 14). The parser must then retrieve the filler from its memory and connect it with the gap. Then, the meaning can be interpreted.

Filler-gap dependencies are also unbounded, meaning that there is no apparent limit to the number of constituents that can separate the filler and the gap. The filler and the gap can be separated by an unlimited number of clauses, as exemplified in (3). The brackets mark the edges of subordinate clauses.

(3) John asked [*which pictures$_i$* Mary said [that Isak thought [that the photographer took __$_i$]]]

By looking at (3), it is clear that long-distance movement is possible. However, as discussed below, this type of movement has some restrictions.

2.1.2 Island effects

Although filler-gap dependencies are present in most languages with no apparent restriction on the linear or structural distance between the displaced phrase and gap, there exist certain guidelines which language users follow unconsciously. Compare the sentences:

(4) a. *What*$_i$ did Dave claim [$_{CP}$ that he saw __$_i$ last night]?

    b. **What*$_i$ did Dave make [$_{DP}$ the claim [$_{CP}$ that he saw __$_i$ last night]]?

Both sentences above involve a wh-word which has moved from its base position. (4a) is an acceptable sentence which involves wh-movement from a complement clause. The filler *what* has moved from within the embedded clause to the front of the sentence for question formation. However, you can see that a similar movement is not allowed in (4b), as indicated by the asterisk. In this example, the filler originates within a complex DP, a domain that blocks movement from the embedded clause. More specifically, a complex DP which is situated inside a CP that is complement to a N. Thus, it appears that some syntactic constituents (or phrases) block filler-gap dependencies. These types of phrases are referred to as *islands* (Ross, 1967). *Island effects* refer to the unacceptability which happens when a gap is situated inside these entities (Sprouse & Hornstein, 2013). The name is meant to be purely iconic, meaning that islands are constituents a phrase cannot escape[1].

Syntactic theories have tried to explain island effects in terms of constraints on A'-movement[2], such as Subjacency (Chomsky, 1973; 1977)[3]. Subjacency proposed that all long-distance movement that appeared to cross more than one clause, was actually broken down into a series of smaller movements. The result is *successive-cyclic movement* of phrases[4], where the phrases stop at certain landing sites before moving up in the sentence. Under Subjacency, island effects occur when a moved phrase is blocked from stopping over at an intermediate landing site.

To illustrate how successive-cyclic movement works, example (3) will be repeated with landing sites marked.

(5) John asked [$_{CP}$ *which pictures*$_i$ [$_{C'}$ Ø [$_{TP}$ Mary said [$_{CP}$ t$_i$ [$_{C'}$ that Isak thought [$_{CP}$ t$_i$ [$_{C'}$ that the photographer took __$_i$]]]

The intermediate landing sites are believed to be specifiers of CP in English, marked by a 't' for trace. Successive-cyclic movement is not possible if the specifier of a CP is occupied by a

---

[1] Due to the boundaries of the DP in (4b), movement from the inner CP is not allowed by grammar. This type of island effect is called a complex DP island, referring to the determiner phrase which forms the borders of the island.

[2] Syntactic movement of a filler to non-argument positions, usually to the specifier position of a CP.

[3] It was earlier believed that Subjacency could explain a significant portion of island effects, at least in English. The Subjacency Condition provides a general account of why extraction from islands are perceived as ungrammatical by referring to how many nodes a phrase can cross on its way to the CP-specifier position in the matrix clause. A problem is that the Subjacency Condition was motivated by some island constraints (complex DP, RC and wh-islands and some subject islands), but not all.

[4] Currently, The Phases Framework (first proposed by Chomsky, 2000) is adopted by the Minimalist Framework to explain the notion of successive cyclic movement (as a universal feature in every natural language).

different phrase. This is exemplified in (6) where the CP specifier position in the relative clause is occupied by the relative pronoun 'who'.

(6)   *What_i did John see [_{CP} [_{TP} the teacher [_{CP} **who** [_{TP} discovered that the students were stealing _____i]

2.1.2.1 Relative Clause Islands and Subject Islands

In this thesis we are interested in two types of islands: **Relative clause (RC) islands and subject islands.** We will start by describing relative clause islands. Relative clauses are embedded clauses that modify nominals (Alexiadou, Law, Meinunger, & Wilder, 2000, p. 2). The clause *that sold hats* in (7) is a RC that modifies the noun *woman*. Although movement can happen within a relative clause, trying to move phrases out of a relative clause results in an unacceptable construction. Neither wh-movement (8) nor relativization (9) out of an RC is allowed in English.

(7)   Lisa saw a woman [_{RC} that sold hats].

(8)   *What_i did Lisa see a woman [_{RC} that sold ____i]?

(9)   *Those_i are the hats_i that Lisa saw a woman [_{RC} that sold ____i]

Similarly, phrases in subject position appear to be islands. The DP *the book about pregnancy* in (10) is a subject.

(10) John thought that [_{SUBJ} the book about pregnancy] was poorly written.

(11) a. John thought that [_{SUBJ} the book about what_i] was poorly written?

b.*What_i did John think that [_{SUBJ} the book about ____i] was poorly written?

In (11a), *pregnancy* has been replaced by the wh-word *what*. In (11b) the wh-word *what* has moved from the DP subject in the subordinate clause to the specifier of CP in the matrix clause. This movement results in an unacceptable sentence. It is widely attested that English speakers are sensitive to both subject islands and RC islands (Sprouse & Hornstein, 2013).

The filler-gap dependency this thesis is particularly interested in consists of a RC located inside a subject phrase: the constituent *the scientist that had come up with a revolutionary theory* in (12). Extraction from this constituent results in an unacceptable sentence, as in (13).

(12) [_{SUBJ} The scientist [_{RC} that had come up with a revolutionary theory]] won a lot of awards for her work.

(13) *That was *the revolutionary theory_i* that [_{SUBJ} the scientist [_{RC} that had come up with ____i]] won a lot of awards for her work.

We will refer to constructions (13) as subject RC islands.

2.1.3 Cross-Linguistic Variation in Island effects

It is generally assumed that island effects reflect innate universal constraints on sentence structure-building (Kush et al., 2018, p. 744). However, this universalist approach would assume that island constraints are similar across all natural languages. Contrary to this initial prediction, there is attested cross-linguistic variation in some island effects. Sprouse and Hornstein (2013, p. 4) report that English demonstrates at least eight different types of island effects. Other languages, such as Italian and Spanish, seem to have fewer restrictions (Rizzi, 1982 for Italian; Torrego, 1984, for Spanish). More importantly, the Scandinavian languages seem to have even fewer restrictions. Notably, Norwegian differs from English in that it appears to allow filler-gap dependencies into some islands, like relative clauses (14).

(14) *De blomstene*$_i$ kjenner jeg   en   mann [$_{RC}$ som   selger ____$_i$.]
     Those flowers know   I   a   man   who sells
     'Those flowers, I know a man who sells.'

(Maling and Zaenen, 1982, p. 232, ex. 4)

In (14) the filler *de blomstene* has moved to the front of the sentence to add emphasis, a type of A'-movement called Topicalization. Topicalization is the movement of a topical constituent to the front of the sentence. Mainland Scandinavian languages are known to be particularly liberal with this movement, especially in speech.[7] [8]

---

[7] Although topicalization is allowed in some cases in English, a movement of this sort is usually not considered grammatical: *Those flowers$_i$ I know a man who sells ___$_i$.

[8] Even though topicalization is frequently used to in Norwegian, there are some cases of topicalization which have been found ungrammatical (i).

(i)   *$Rødsprit_i$ slipper vi   ingen [$_{RC}$ som har drukket ___$_i$] inn.
            Red.spirit   let   we   nobody that has drunk   in
                'Red spirit, we let nobody in that has drunk (that)'

(Taraldsen, 1982, p. 206, ex. 9)

A bigger puzzle is that the phrase seems to be accepted by Norwegian speakers when the relative clause is in a sentence-final position (Taraldsen, 1982, p. 206).

(ii) $Rødsprit_i$ slipper vi ingen inn [$_{RC}$ som har drukket ___$_i$].
            Red.spirit   let   we nobody in that has drunk

It is therefore hard to say for certain that relative clauses are not islands in Norwegian. In a recent study, Norwegian speakers were sensitive to the extraction of a wh-word in a relative clause (Kush et al., 2018). A possible explanation for the results is that the participants were not given a hypothetical discourse context where the experimental sentences were plausible. Engdahl (1997) argues that island effects can be affected by discourse context. In a later experiment involving topicalization in different types of islands, Kush, Lohndal & Sprouse (2019) found that Norwegian speakers were more likely to judge sentences more acceptable if they were given a context where the participants could imagine the sentences being uttered.

6

Generally, subject phrases are considered islands in the Scandinavian languages as well. This is exemplified in (15) where the filler is extracted from within the subject phrase, and (15) where the filler has moved from a sentential subject phrase.

(15) a. [SUBJ Lingvistene   på det foredraget] skal danse på scenen.
            The linguists at that lecture   will   dance on the stage
            'The linguists at that lecture will dance on stage'

      b. *Det foredraget*i* skal [SUBJ lingvistene   på ____i] danse på scenen
                that lecture   will   the linguists at   dance on the stage

(16) a. [SUBJ Påstanden om   at hun gjorde det] har aldri vært sann.
            the claim about   that she did   it   has never been true.
            'The claim that she did it has never been true'

      b. *Det*i* har [SUBJ påstanden om   at   hun gjorde __] aldri   vært sann.
          that has   the claim   about what she did   never   been true

As for subject RC-islands, they are also generally believed to be unacceptable by Norwegian speakers.

(17) [SUBJ Jenta [RC som nettopp hadde kjøpt   seg   nye ovnsvotter]] likte
            the girl   that   just   had bought herself new oven.mittens liked

      å lage   mat.
      to make food

      'The girl that just bought new oven mittens (for herself) liked to cook'.

(18) *Det   var   *de nye ovnsvottene*i* som [SUBJ jenta [RC som nettopp hadde kjøpt
        those were the new oven.mittens that   the girl   that   just   had bought

      seg __i]] likte å lage mat.
      herself   liked to make food

There are, however, ways subject RC-islands can become acceptable in Norwegian, but not in English.

2.1.4 Parasitic gaps

Although it was stated earlier in the text that gaps cannot occur inside islands, this seems to not always be the case. In the last decades, research on island constraints have been particularly focused on a phenomenon referred to as *parasitic gap constructions*. A parasitic gap construction occurs when an illicit gap inside of an island becomes acceptable due to

the presence of an additional licit gap located outside the island (Engdahl, 1983, p. 5), exemplified in the sentence below:

(19) This is the kind of food$_i$ [you must cook ___$_i$ [before you eat ___$_{pg}$]].

$$\text{(Engdahl, 1983, p. 5, ex. 2)}$$

The illicit gap is marked '*pg*' and the licit gap is with '*i*'. Most English speakers will accept (19) without hesitating, ignoring the fact that a gap is located inside an adjunct island. The fact that the acceptability of (19) depends on the presence of a licit gap in the object position of "cook" is demonstrated by the unacceptability of (20):

(20) *This is the kind of food$_i$ you must cook rice [before you eat ___$_i$].

Parasitic gaps are also acceptable in some subject islands. Consider (21), where a simple gap after "cover up" is unacceptable (a), but the same gap becomes acceptable if it is parasitic (b).

(21) a. *Those were *the secrets$_i$* that [the attempts to cover up ___$_i$] ultimately revealed the problems to the public.

b. Those were *the secrets*$_i$ that [the attempts to cover up ___$_{PG}$] ultimately revealed ___$_i$ to the public.

Although constructions like (21b) are rather uncommon in English, they can be accepted. The parasitic gap seems to be acceptable inside a complex subject when the clause is infinite like in (21b). Phillips (2006) reported that native English speakers accepted parasitic gap configurations in both online and offline experiments in similar infinitival subject sentences. Interestingly, a parasitic gap inside a subject is not deemed acceptable by native speakers if the parasitic gap occurs inside a *finite* relative clause within the subject, such as (22).

(22) *Those were *the secrets$_i$* that [that the politician attempted to cover up ___$_{pg}$] ultimately revealed ___$_i$ to the public.

Not surprisingly, acceptance of parasitic gaps also varies cross-linguistically. The Scandinavian languages appear to have fewer restrictions on parasitic gaps, when seen in comparison with English (e.g. Engdahl, 1983; Christensen & Nyvad, 2011). Specifically, Swedish has been reported to allow parasitic gaps inside *finite* RCs inside subject phrases (Engdahl, 1983)[9]:

---

[9] Individual preferences among language speakers have been reported in regards to parasitic gaps. Engdahl (1983) points to individual variations between the Swedish speakers in her experiment, noting that 'some speakers are very restrictive about which positions they do accept parasitic gaps in, others are more permissive' (p. 8).

(23) a.Räkna upp *de filmer$_i$* [$_{RC}$ som alla   [$_{RC}$ som har sett ____$_{pg}$] tyckte bra om ___$_i$.]
        List   those films   that everyone who has seen   liked   a lot

'List the films that everyone that has seen them and liked a lot.

   b. Kalle är *en kille$_i$* [$_{RC}$ som ingen [$_{RC}$ som träffet ____$_{pg}$] kan tåla ___$_i$.]
      Kalle is a guy   who no one   who (has) met   can stand

'Kalle is a guy that no one who has men him can stand.

Although the example provided is in Swedish, it is likely that the same sentences in Norwegian will be considered grammatical by native speakers.

(24) a. Rams opp *de filmene$_i$* [$_{RC}$ som alle   [$_{RC}$ som har sett __$_{PG}$] syntes __$_i$ var bra].
         List   those films   that everyone   who has seen   liked   a lot

   b. Kalle er *en gutt$_i$* [$_{RC}$ som ingen [$_{RC}$ som har truffet____$_{PG}$] takler____$_i$]*.*
      Kalle is a guy   who no one   who has met   stands

If a speaker expects a parasitic gap when processing a filler-gap dependency, they might anticipate gaps inside islands if there is a chance of it being salvaged later in the sentence. This can be the case with the ungrammatical subject RC islands (example 20) listed in chapter 2.1.3. By replacing the complement of the verb "å lage mat" with a licit gap, the sentence becomes less unacceptable (and arguably acceptable).

(25) ?Det   var   *de   nye ovnsvottene$_i$* som [$_{SUBJ}$ jenta [$_{RC}$ som nettopp hadde kjøpt
       those were the new oven.mittens that   the girl   that   just   had   bought

    seg ___$_{PG}$]] likte ___$_i$.
    herself   liked

'Those were the new oven mittens that the girl that just had bought them liked'

If Norwegian speakers accept constructions such as (25), it might have some implications for L2 processing. This question will be discussed in the following chapters.

2.2 Parsing filler-gap dependencies

2.2.1 Active Gap-Filling Strategy in L1

There are several theories that attempt to explain the strategies a parser can use when encountering a filler. Two possible strategies will be presented: a *passive* and an *active gap-filling strategy (AGF)* (Clifton & Frazier, 1989; Fodor, 1979). The first strategy involves making no predictions regarding where the gap site is located. Before deciding where the gap site is, the parser therefore reads the sentence, then looks for unambiguous evidence for the gap in the sentence. In (26) there is only one possible gap site when the sentence is

fully parsed. However, there is an uncertainty about where the gap is located in the sentence during incremental left-to-right processing.

(26) *Which girl*~B~ do you believe ____~A~Timmy likes _____~B~ a lot?

In this sentence, a left-to-right parser may consider a dependency at the apparent gap (A) before it has seen the rest of the sentence, as the filler *which girl* is a possible object of the verb *believe*. A parser might consider a gap position at (A) since it does not know that another verb *likes*, which hosts the actual gap site (B), is coming later in the sentence. A passive parser would process the sentence by not positing a gap at the first possible gap site; it would read the sentence until it gets sufficient evidence for the actual gap position, before deciding that the gap is located at B. Thus, a passive parser will not posit a gap until clear evidence for the gap position is presented.

The alternative strategy states that the parser actively searches for gaps after encountering a filler and predictively posits gaps before getting unambiguous evidence that the gap location is correct (Fodor, 1978; Crain & Fodor, 1985; Clifton & Frazier, 1989; Traxler & Pickering, 1996). This active search for gap sites can lead the parser to initially make an inaccurate structural analysis while reading a sentence. When the parser receives evidence that the analysis is incorrect, it then needs to reanalyze which results in a higher processing load. In online experiments, this is often shown as an increased reading time. Looking back to example (26) an active parser would try to posit gap (A) after seeing the verb *believe*. However, after seeing the noun phrase *Timmy*, the parser would know that the initial structural analysis was wrong. It would then need to reanalyze the whole sentence by starting to look for a new gap. When the parser reaches the verb *likes*, it realizes that the gap is located here (B).

There are several experiments in the literature that have investigated active gap-filling. For example, Stowe (1986) looked into native English participants in a self-paced reading experiment on filler-gap dependencies. Participants in the experiment were asked to read sentences like (27) which came in two variants: In (27a) there was a filler *who* that introduced the embedded clause. In (27b) there was no filler; instead the conditional *if* introduced the embedded clause. As there was no filler, there was no gap in the embedded clause that the participants could fill in.

(27) a. My brother wanted to know *who*~i~ Ruth will bring us home to __~i~ at Christmas

   b. My brother wanted to know *if* Ruth will bring us home to Mom at Christmas.

The participants in this study displayed a longer reading time for *us* in (27a) than in (27b). The slowdown at *us* was interpreted as participants expecting a gap for the filler *who* after reaching the transitive verb *bring*. As the participants expected a gap in this position, seeing *us* in this position was surprising, since *bring* cannot take two different direct objects. Thus, participants had to erase the predicted dependency they tried to create between *who* and *bring* (i.e. My brother wanted to know who~i~ Ruth will bring __~i~) before they continued to parse the sentence. In (29b) *who* has been replaced with the conditional *if,* which does not need to find a gap. Therefore, no gap will be predicted after *bring* and no reanalysis will be required when *us* is encountered.

10

It is possible to distinguish an AGF strategy from a passive strategy in sentences with temporary ambiguity, such as (27a). A study which is often cited in regards to active gap filling is Traxler & Picker (1996). Traxler and Pickering recorded participants' eye movements when they were exposed to filler-gap dependencies in sentences with temporary incremental ambiguity. These types of sentences seem to be structurally ambiguous up until a point in which the actual and only interpretation becomes clear. When reading through the sentences in (28) from left to right, the first verb the parser encounters is *wrote*. *Wrote* is optionally transitive, meaning that the verb *can* take an argument (but it is not obligatory). It is expected that an active parser will try to interpret the filler as the direct object of *wrote*, as it is a possible argument of the verb. However, the actual gap appears later in the sentence.

(28) a. We like *the book$_i$* that the author wrote __$_X$ unceasingly and with great dedication about__$_i$ while waiting for a contract.

b. We like *the city$_i$* that the author wrote __$_X$ unceasingly and with great dedication about __$_i$ while waiting for a contract.

The '__' are added to show the 'possible' gap sites where an incremental parser might posit a gap in the sentences. The false gap site "__$_X$" is located after the first optionally transitive verb *wrote*, and the actual gap is located after the PP complement *about*. Traxler and Pickering used plausibility manipulation to elicit evidence for active gap-filling; the filler varies between two DPs, *the book* and *the city*. *The book* fits semantically as a plausible object of the optionally transitive verb *wrote*, whereas the other filler *the city* does not. Participants in Traxler and Pickering's study exhibited an increased eye-gaze duration at the optionally transitive verb (wrote) when reading the implausible (28b), but not (28a), demonstrating a 'filled-gap effect'.

The filled-gap effect in the plausible condition suggests that the participants tried to form a dependency at the earliest possible gap location. In comparison, the effect was not present in the implausible constructions as the filler was not a semantic fit with the critical verb. This study gave evidence of a preference for active gap filling in (L1) sentence processing. Even though a strong bias for creating the shortest possible dependency is attest in parser, would that imply that parsers violate grammatical constraints?

2.2.2 Island Sensitivity in L1 Parsing

As discussed above, most studies suggest that parsers attempt to form the shortest possible filler-gap dependencies. As active-gap filling seemed to be the go-to method for language processing, Traxler & Pickering (1996) were interested in whether parsers went for an active gap-filling strategy *regardless* of grammatical constraints. In particular, whether parsers were sensitive to island constraints. Thus, they added an additional factor in their experiment: islandhood. The islandhood factor was added as an additional relative clause inside the subject phrase in (28). The first "fake" gap site is now situated inside an island.

As fillers cannot move out from islands, a gap position inside one should not be possible[10]. In (29) it is not syntactically possible for the filler (*book/city*) to be linked with *wrote* because the verb is now inside a RC, which is itself inside a subject.

(29) a. We like *the book*$_i$ that [$_{SUBJ}$ the author [$_{RC}$ who wrote ___$_x$ unceasingly and with great dedication]] saw ___$_i$ while waiting for a contract.

b. We like *the city*$_i$ that [$_{SUBJ}$ the author [$_{RC}$ who wrote ___$_x$ unceasingly and with great dedication]] saw ___$_i$ while waiting for a contract.

Traxler and Pickering compared the reading times in the island conditions (29) with the non-island conditions (28). If there was a filled-gap effect at the critical verb in both the island and non-island implausible constructions, it would indicate that the parsers would form dependencies even when the dependency would ultimately be ungrammatical. The results showed that there was no significant increase in eye-gaze duration in the plausible, island condition (29a) when compared to the implausible, island condition (29b). However, as we saw in the section above, the participants did show a plausibility mismatch effect in the non-island conditions. These findings suggest that the native speakers are able to constrain their active gap filling strategy in situations where grammar does not allow it. Other studies have also reported similar results (e.g. Stowe, 1986; McKinnon & Osterhout, 1996). These studies show that English readers follow grammatical constraints when parsing language in their L1.

2.3 L2 Processing

Active-gap filling is one of the L1 processing strategies which have also been evident in L2 speakers (Williams, Möbius & Kim, 2001; Omaki & Schulz, 2011; Kim et al., 2015). A focal question in this thesis is whether the L2 users are able to constrain this strategy when faced with illicit gaps during online processing. This chapter will start by talking about L2 processing in general, before introducing the main theories this thesis will discuss.

2.3.1 General overview

Research on second language acquisition and processing attempts, in part, to determine to what extent L2 users have the same abstract representations as native speakers (Mackey & Gass, 2015, p. 58). As differences between L1 and L2 processing have been noted, researchers have put forward different theories to explain the observed differences.

Many theories of L2 processing assume that non-native parsing is essentially the same as L1 processing. It may be less automatized and may operate more slowly, but the sentence representations that the L2 processor creates are essentially similar to

---

[10] Unless we are dealing with a parasitic gap, discussed in chapter 2.1.4.

representations created by a native parser. It has been argued that the increased processing load an L2 speaker experiences makes it difficult to maintain the filler in memory until the gap position is revealed (e.g., Cunnings, 2017). According to this claim, keeping a filler in memory until you encounter the actual gap is harder for non-native speakers, as L2 processing is cognitively more demanding. Hopp (2011) argues that L1 and L2 processing are qualitatively similar to each other, but that individual variations in working memory, which also occur in native processing, may be a key factor in the observed processing differences.

There are, however, reasons to believe that L2 users pursue qualitatively different parsing strategies in their target language. Although researchers agree that L2 processing is in general slower than native language processing, some studies involving L2 users indicate that L2 users in some cases pursue *non-nativelike* parsing strategies. Several theories claim that a L2 speaker will most likely never reach the same automatized level of fluency which a native speaker has (Clahsen & Felser, 2006; Paradis, 2009). Based on this assumption, a nonnative speaker will not be able to utilize all linguistic information available, to the same extent that a native speaker is capable of.

Some of the theories of L2 acquisition and processing that posit deep differences between L1 and L2 processing draw on neurocognitive models of the brain which depict where language is stored (e.g. Ullman, 2001; Paradis, 2009). It is argued that the learning, representation and processing of lexical items and grammar are dependent on two brain memory systems: the *declarative* and *procedural* memory. The models differ in their assumptions on where the declarative and procedural memory is located in the brain. The declarative long-term memory is believed to subserve the conscious learning of facts (semantic knowledge) and concepts, such as lexical items (e.g. words) in one's native language. The procedural memory stores knowledge that is carried out unconsciously, such as automatized actions (motor and cognitive skills) and L1 grammar and syntax. According to Ullman (2001) and Paradis (2009) early L2 learning mainly takes place in the declarative memory. Thus, even the grammatical features in the L2 are accessed consciously, which may explain the less-successful processing patterns.

In addition, it is posited that the usage of the procedural memory tends to decline with age, as a shift of dependence from the procedural memory to the declarative memory increases with age (Ullman, 2001, p. 110). This is relevant for L2 learning in general, as the age one learns a second language seem to be a relevant factor for L2 proficiency (L2 learned as a child vs. late-learned L2). However, this does not necessarily imply that L2 learners can never acquire L2 grammar in their procedural memory; Ullman (2005) predicts that proceduralization of grammatical rules in the target language can happen through sufficient exposure and proficiency (p. 151)[11]. However, the idea that L2 learners depend more on their declarative memory, is still the general consensus.

The D/P model is relevant to this study as we are interested in how L2 learners parse language in real-time. It is believed that online studies, such as self-paced reading studies, are able to tap into one's implicit knowledge of grammar (i.e. they will not leave you time to think about your decisions). If an L2 user has explicit knowledge of a grammatical feature in the target language, but fails to use this knowledge during online sentence processing, it

---

[11] Paradis (2009), on the other hand, states that this is very rarely the case; most L2 learners will never reach this level of fluency.

would suggest that the parser is not being able to utilize all linguistic information and cues during parsing. It is therefore relevant to keep this in mind when conducting studies with L2 users. In the following sections, we will look into a claim regarding L2 users showing a preference for semantic cues over syntactic. Lastly, we will discuss how the native language may also play a role in L2 sentence processing.

2.3.2 Shallow Structure Hypothesis

A native speaker relies on most linguistic information during sentence processing, including syntactic, lexical-semantic, contextual and prosodic cues (Clahsen & Felser, 2006, p. 4). However, Clahsen & Felser (2006) have argued that L2 users may need to prioritize one kind of information at the expense of the others, as they need to be more efficient due to limited memory and processing resources. In other words, differences in native and nonnative online language processing can be explained by L2 speakers producing 'shallower and less detailed' syntactic representations (p. 1). This claim is referred to as the *Shallow Structure Hypothesis (*SSH), stated to be a general property of the L2 user, regardless of how closely related one's L1 and L2 are.

Clahsen and Felser base their theory on studies done on children, adults and L2 processing. In particular, Marinis, Roberts, Felser and Clahsen (2005)'s study which compared native English speakers and advanced L1 Chinese, Japanese, German and Greek learners of English was important for developing their hypothesis. The participants in this study read sentences with long-distance wh-dependencies in a self-paced reading task. In (30) the DP (*the nurse*) is followed by a relative clause introduced by a wh-pronoun (*who*), which functions as the object of the embedded verb (*had angered*). The intermediate verb (*argued*) in (a) permits wh-extraction, but the verb is swapped with the DP *argument* in (b) Thus, there is no intermediate gap site in (b). [12]

(30) a.The nurse *who*$_i$ the doctor argued ___$_A$ that the rude patient had angered ___$_B$ is refusing to work late.

b. The nurse who$_i$ the doctor's argument about the rude patient had angered ___$_i$ is refusing to work late.

The native control group showed reading-time evidence for a pre-gap reactivation of the filler*who* when reaching the position marked (A) in (30a) before reaching the actual gap position at (B). Transformational theories assume that the filler has stopped at the landing site (A) due to the position of a clause boundary, and then it creates an intermediate copy of itself before moving to the front of the sentence (Chomsky, 1995).
Studies have revealed that native speakers show an increased reading time at intermediate gap sites (e.g. Gibson & Warren, 2004). The native speakers in Marinis et al. behaved similarly, in that they slowed down after *argued* in (30a) but not (30b), as there was no intermediate gap site in this sentence. However, the nonnative speakers in this experiment did not show any indications for a reactivation of the filler. Clahsen & Felser argued that this

---

[12] The original experiment had a 2x2 design with extraction/non-extraction and VP/NP factors. For simplicities sake, we will only refer to the extraction sentences here.

finding of a non-native like processing pattern indicated that the L2 users were underusing syntactic information during online language processing.

However, the study above (and the SSH in general) has been criticized. Omaki & Schulz (2011) report that the study did not assess whether the participants had the prerequisite knowledge needed to demonstrate the expected processing behavior (pp. 567-568). Although Marinis et al. checked the participants' proficiency, they did not assess whether the nonnative had the required grammatical knowledge for these types of sentences. The non-native like parsing may be a result from the participants not having acquired the necessary syntactic knowledge in their target language and relied on structural information from their L1 instead.

Other explanations for the findings in Marinis et al. could be due to a different memory architecture in L2 parsers, or simply that the research method was not optimal (Omaki & Schulz, 2011, p. 567). One example being there was no evidence that the L2 parsers had the relevant grammatical knowledge to demonstrate the expected behaviors. To counter these factors, Omaki and Schulz designed a new experiment on filler-gap dependencies and islands using L1 Spanish-L2 English speakers. Given that islands are considered to be purely syntactic constituents, the SSH would suggest that L2 speakers are likely to value lexical-semantic cues during sentence parsing over an island constraint if these were competing. The experimental sentences in Omaki and Schulz were based on the items in Traxler & Pickering (1996):

(31) a. *The book/the city$_i$* that the author wrote ___$_x$ regularly about___$_i$ was named for an explorer.

b. *The book/the city$_i$* that the author [$_{RC}$ who wrote ___$_x$ regularly] saw ___$_i$ was named for an explorer.

Omaki and Schulz wanted to test whether the nonnative speakers were (i) active gap fillers and (ii) sensitive to island constraints despite a preference for immediate gap creation. (31a) involves the same plausibility manipulation used in Traxler & Pickering, and (31b) has an added relative clause. The relative clause is now situated inside an island meaning that a phrase cannot have been extracted from it to form a dependency.

Omaki & Schulz (2011) found that the L2 speakers parsed the sentences similarly to the native English control group. They found a plausibility-mismatch effect at the verb *wrote* in (31a) indicating that L2 participants were active gap-fillers. Additionally, even though the filler *the book* fits plausibly with the 'fake' gap site (after the verb *wrote*), the L2 users did not show any indication of trying to form a filler gap dependency inside the island in (31b). If the L2 users were only driven by lexical-semantic cues, they should have ignored the syntactic island constraint. Instead, they seem to have enough syntactic fidelity to distinguish between islands and non-islands (see also Witzel, Witzel, & Nicol, 2012).

2.3.3 L1 Transfer in Parsing

A major difference in L1 and L2 acquisition, is that the latter involves acquiring a new language when the person is already equipped with a fully developed language system. L1

effects on the target language, or *transfer*, is a fact in early L2 development. However, exactly what linguistic information transfers from the L1 and how this happens, is still unclear. Researchers are also not sure if transfer effects are still present in fluent L2 processing. There seems to be inconclusive evidence on this matter, as some studies show evidence for processing transfer in second language comprehension (Juffs, 1998; Kim, Baek & Tremblay, 2015), while others do not (Marinis et al., 2005, Cunnings et al., 2010). The optional view is that there are some underlying principles of non-native processing in general, such as the SSH, instead of transfer from one's L1. This thesis is interested in whether the syntax in one's native language is activated during second language processing, and will consider the probability for this theory.

The transfer of syntactic properties of one's native language has been evident in studies (Juffs, 1998; Kim, et al. 2015). In an online study, Kim et al. (2015) studied Korean and Spanish speakers' processing of English island constraints in wh-dependencies using a stop-making-sense task. These groups were chosen as Spanish uses overt movement, similarly to English, to form wh-dependencies. Additionally, Spanish and English display many of the same island effects, including the RC island effect. Korean, on the other hand, is a wh-in-situ language, meaning that the wh-phrase does not need to move from its base position in simple wh-questions. More importantly, Korean does not display RC island effects in situ wh-questions (Kim et al., 2015, p. 386). Therefore, we might expect the Koreans to process English wh-dependencies differently than the Spanish speakers.

Kim and colleagues' study involved experimental sentences similar to the one below, based on Traxler & Pickering (1996):

(32) a. I wonder *which book/city$_i$* the author wrote __$_x$ passionately about __$_i$ while he was travelling.

b. I wonder which book/city the author [$_{RC}$ who wrote __$_i$ passionately] saw __$_i$ while he was travelling.

(32b) consists of a wh-dependency across a relative clause which the filler has moved around. Similar to Omaki & Schulz (2011), there is an optionally transitive verb inside the RC which the filler could plausibly be the object of in one of the conditions. A comparison of the island (32b) and non-island (32a) conditions revealed whether the relative clause island constraint guided nonnative processing, and a comparison of plausibility/implausibility in (32a) showed whether the participants were active gap fillers. Kim and colleagues were interested (like Omaki and Schulz) in whether the L2 participants would show signs of active gap filling in (32a), but not in (32b).

The Stop-Making Sense task gathered the participants' reading times per region and implausibility detection rates (i.e. the participants pressed a button when the sentence did not make sense anymore). The task therefore provided online reading times and offline judgments of plausibility. The results showed that the Spanish speakers displayed a similar processing pattern as their native English control group. In (32a) the Spanish speakers pressed the button more in the implausible condition than in the plausible condition after the intermediate verb*,* indicating that they tried to link the filler (which book/city) with the verb *wrote*. The reading times display the same pattern (i.e. slow-down in the implausible conditions after the critical verb, but not in the plausible conditions). The same effect was

not there in the island conditions (32b) for the Spanish (and English) speakers. The results from the (a)-sentences indicate that the Spanish and English speakers were active gap fillers, and the results from the (b)-sentences suggest that they were able to constrain their active search for gaps when the critical verb was inside an island domain.

The Korean participants, on the other hand, behaved slightly differently. The results (both reading times and detection rates) from the non-island sentences indicated that they were active gap fillers, as their results replicated the Spanish and English groups' behavior after the critical verb. Similarly, the plausibility detection rates indicated that they respected the island constraints in (32b), as there was no difference in judgments between the plausible and implausible condition.

However, the reading times showed that the Korean group displayed a plausibility mismatch effect at the intermediate verb in the island sentences. This suggests that they initially tried to automatically fill a gap inside the RC, in violation of island constraints. Kim et al. theorized this to be due to them having explicit knowledge of English island constraints, as speakers are able to consult with their explicit knowledge of grammar during offline judgments. This finding suggests that the Korean speakers tried to posit a gap inside the RC unconsciously, even though they were aware of it being ungrammatical.

The results in Kim et al. supports the theory of transfer because Korean does not exhibit the relative clause island effect. As relative clauses are not islands in Korean, it might be the case that the Korean speakers transferred this property when processing English filler-gap dependencies with island domains.

2.4 The Current Study

The findings from the studies mentioned so far bring us to why we wanted to investigate Norwegian speakers' behavior during processing of filler-gap dependencies. There are two ways pointed out in previous chapters in which Norwegian differs from English: (i) Extraction from relative clauses in Norwegian seems acceptable and (ii) parasitic gaps inside finite subject RCs are arguably acceptable in Norwegian.

If Norwegians allow filler-gap dependencies into relative clauses and finite subjects, we can expect them to allow this 'violation' during online English sentence parsing. Then, if there is transfer of L1 parsing strategies to L2, we can expect the Norwegians to pursue active-gap filling inside subject and relative clause islands, where native speakers would not. In essence, they would act similarly to the Korean speakers in Kim et al., and not like the Spanish speakers with similar syntactic structure in their native language.

3 Method

3.1 Participants

Fifty-seven native Norwegian speakers participated in the experiment. The participants were recruited through public posts on Facebook and Innsida (a site for students and employees at the Norwegian University of Science and Technology, NTNU), or through an undergraduate class at NTNU. Six participants were excluded due to low accuracy in the comprehension questions during the self-paced reading task. The participants had a mean age of 24.3 (SD = 3.9) with 36 females. Participants provided their age, whether their native language was Norwegian whether they had spoken any other languages as a child (*yes/no*). Further, they self-rated their proficiency in English, while also estimated their average exposure to English on a regular basis. This included average media exposure per day and average time spent using English per week. All participants took part voluntarily, and no identifying data was recorded.

3.1.2 The participants' proficiency level

The experiment was designed similarly to the self-paced reading experiment conducted by Omaki & Schulz (2012), with the exception of a method for testing participants' English proficiency. Omaki & Schulz used a 'C-test', a gap-filling test based on the reduced redundancy principle (Eckes & Grotjahn, 2006, p. 291). In gap-filling tests, participants are exposed to sentences where parts of words are deleted. The participants task is to fill in the missing parts to gather the meaning of the sentence. The participants' performance is used as an indicator of general language proficiency.

In our experiment, the Norwegians self-reported their proficiency. As Norwegians are known to be highly proficient in English due to frequent exposure through school and media, it is assumed that the participants' proficiency is somewhat high (Simensen, 2009). The participants have most likely finished the mandatory English instruction in primary and secondary school in the Norwegian education system, which ranges between 10-13 years (approximately 728 hours) depending on the participants' age and electives (The Norwegian Directorate for Education and Training, 1997; 2013). On a 7-point scale, all participants rated their proficiency level between 4 and 7 (M = 5.51, SD = 0.75). The participants also had to report how much English media they consumed on average each day. The answers were converted into a numerical scale from 1-4 (1 = <1 hour, 4 = 5+ hours a day). The participants reported a high average media exposure per day (M = 3.35, SD = 0.68). The participants were also asked to report how often they used English per week. The answers were again converted into a scale from 1-6 (1 = 0-1 hours a week, 6 = 20 + hours a week). A mean of 4.10 (SD = 1.62) suggests a relatively high English usage per week.

More importantly, the participants had to state when they started learning English. Most participants started learning English when they were between 5-7 years old (M = 6.25, SD = 2.13), well before late puberty, which has recently been cited as the end of the critical period of second language acquisition (Hartshorne, Tenenbaum, & Pinker, 2018). Although a critical age for L2 acquisition is debated, it is assumed that early L2 acquisition, in addition

to frequent exposure to English, is enough to conclude that the participants are highly proficient in the target language. The obtained information is summarized in Table 3.1.

To check whether participants were aware of island constraints in English, the participants were asked to complete an acceptability judgment task after the self-paced reading procedure.

Table 3.1

Demographic information of the participants. The parentheses indicate the scale used to measure the participants' reported responses.[13]

Question	Mean	SD
Participant's age	24.29	3.91
Age of first instruction	6.25	2.13
Reported English proficiency (1-7)	5.51	0.75
Avg. English media exposure /day (1-4)	3.35	0.68
Avg. English usage /week (1-6)	4.10	1.62

3.2 Experimental design

3.2.1 Overall procedure

The experiment consisted of two parts: (a) a self-paced reading task and (b) an acceptability judgment task. These tasks were implemented on the online experimental platform Ibex Farm (Drummond, 2012). The estimated time of completion for all tasks was around 30 minutes. The participants entered the experiment via a link sent to them on a social media platform or e-mail. The experiment was conducted on the participants' personal computer to ensure we would be able to recruit enough participants. As the data collection happened without controlled supervision, it is more likely that outside factors could have influenced the data (e.g. a noisy environment). The goal was therefore to make sure that we had a sufficient number of participants to combat this issue. The instructions for each task were written in English.

This chapter is organized after the order in which the items appeared in the experiment: Firstly, the materials and procedure used in the self-paced reading task will be

---

[13] A table of the participants' yes/no responses to certain demographic/background questions can be found in the appendices.

presented. This will be followed by an explanation of the materials and procedure used in the judgment task.

3.2.2 Self-Paced Reading Task

Self-paced reading was chosen as a suitable task for this thesis as it is relatively easy, cheap, and timesaving to implement, compared to its more advanced counterpart eye-tracking (Kaiser, 2013, p. 141). More importantly, self-paced reading can be used to uncover processing difficulty in a specific part of a sentence. This experimental design is based on the notion that the eyes are windows on cognition, meaning that the amount of time it takes to read a word reflects the amount of time a person needs to process it (Jegerski, 2014, p. 23). Processing difficulty is evident through increased reading time (RT). Reading time studies can give us information about many different types of processing difficulties, or where processing is more effortful than in other places of the sentence.

It is believed that online experiments will not let participants rely too much on their explicit knowledge of language (Jegerski, 2014, p. 28), which is important in this thesis as we are not concerned with this area of research. Instead, this thesis investigates whether L2 users can rely on syntactic cues during automatic sentence processing, and whether or not the L2 users are using a native-like processing strategy.

3.2.2.1 Materials

The self-paced reading task consisted of 24 target items like (1), similar to the sentences used in Omaki & Schulz (2011)[14]. Items followed a 2x2 factorial design that crossed two factors: *plausibility* and *islandhood*.

(33) a. Nonisland, implausible
The city [$_{RC}$ that [$_{SUBJ}$ the author] wrote regularly about]] was named after an explorer.

b. Nonisland, plausible
The book [$_{RC}$ that [$_{SUBJ}$ the author] wrote regularly about] was named after an explorer.

c. Island, implausible
The city [$_{RC}$ that [$_{SUBJ}$ the author [$_{RC}$ who wrote regularly]] saw] was named after an explorer.

d. Island, plausible
The book [$_{RC}$ that [$_{SUBJ}$ the author [$_{RC}$ who wrote regularly]] saw] was named after an explorer.

The factor *plausibility* controlled whether the filler DP (*the city/book*) was a semantically plausible or implausible object of the first verb in the sentence (*wrote*). Parsers might first posit a gap site in (a) and (b) after *wrote*, as it is an optionally transitive verb which can

---

[14] The items for both tasks (SPR and the acceptability judgment task) can be found in the appendices.

take an object. Ultimately, the actual gap site is revealed later in the sentences. In (a/b), the gap comes after the preposition *about*. The design included an adverb (*regularly*) after the critical verb to give room for possible spill-over effects. A spill-over region is useful as possible effects may not show up at the region of interest; it may be delayed.

The island sentences differ from the non-island sentences by having an additional relative clause embedded inside the subject phrase. In island sentences, the optionally transitive verb (*wrote*) is located inside this relative clause. The filler (the city/book) should not be interpreted as the object of the optionally transitive verb (wrote) in these conditions due to island constraints.

Each experimental sentence consisted of 8-11 regions. The island sentences had an extra region in which the complementizer (*who*) was introduced. With the exception of the extra region in the island conditions, the experimental sentences had the same number of regions up until the end of the relative clause (i.e. where the filler was retrieved). The 24 items were counterbalanced on four lists. The participants saw a sentence from each item only once during the experiment, following a Latin Square Design (Stowe & Kaan, 2006, p. 49). Additionally, 44 filler items of similar length and complexity were added to the list of items participants were exposed to during the experiment.

3.2.2.2 Procedure

The experimental sentences were presented phrase-by-phrase. The phrases appeared in linear succession and in a noncumulative fashion (Jegerski, 2014). The experimental sentences were more suited to use phrase-by-phrase segmentation than word-by-word segmentation, as the sentences were particularly long (8-11 regions). The phrases generally involved 1-2 words, and phrases were made by only grouping together a determiner with its noun, and prepositional phrases together. The critical sentences were preceded by four practice sentences and a set of instructions.

The participants used the keyboard during the experiment to respond to the stimuli. 'SPACE' was pressed to move forward from each region in the self-paced reading task, and 'D' and 'K' were used as 'YES'/'NO' respectively for the comprehension questions. The participants received feedback after the questions if they were answered incorrectly. Additionally, the participants were told to take any necessary breaks when the prompted accuracy questions were on the screen so as not to interfere with the reading time measure.

3.2.3 Acceptability Judgment Task

A 7-point acceptability judgment task was used to assess the participants' knowledge of English island constraints. By combining SPR with an off-line method such as a judgment task, we are able to get more data on the participants' interpretation and proficiency in the target language when combined with an online method (Kaiser, 2013, p. 137). It was necessary for the participants to have knowledge of English RC and subject island constraints to interpret their behavior in the SPR. If they did not have any intuitions that island violations in English were unacceptable, there is no reason for them not to form a dependency in the island sentences (35c)-(35d) in the SPR task.

3.2.3.1 Materials

Four different comparisons were tested in the acceptability judgment task: (i) parasitic gaps inside finite subject RCs; (ii) parasitic gaps inside infinitival subjects; (iii) gaps inside subject RC islands; (iv) gaps inside RC-complements.

The parasitic gap constructions (i)-(ii) were tested to rule out the possibility of participants expecting a licit gap outside the relative clause. In Phillips (2006) it was shown that a parser is able to anticipate a gap inside certain islands if it can be made grammatical by an upcoming grammatical gap. If the participants were not aware that gaps inside finite RC islands cannot be salvaged, they are predicted to proceed with their active gap filling strategy inside the RC island. Judgements of grammatical English infinitival parasitic gap constructions (ii) were added for comparison.

Items (iii) and (iv) were added to assess the participants' knowledge of the English RC island constraints. Specifically (iii) included sentences with subject RC island constraints, and (iv) included sentences with an RC island constraint. (iv) is grammatical in Norwegian, and was added to ensure there were no transfer effects from Norwegian. The experimental sentences will be listed below, where (a) denotes the grammatical conditions, and (b) the ungrammatical conditions.

(34)

i. Parasitic gaps inside infinitival subjects
a. Grammatical sentence
That was the university$_i$ that [$_{RC}$ the woman who had donated money to a charity] had studied at ___ $_i$ for her law degree.

b. Ungrammatical sentence
That was the university$_i$ that [the woman who had donated money to ___$_{pg}$] had studied at ___$_i$ for her law degree.

ii. Parasitic gaps inside infinitival subjects
c. Grammatical sentence
Those were the secrets$_i$ that [$_{RC}$ the attempts to cover up ___$_{pg}$] ultimately revealed ___$_i$ to the public.

d. Ungrammatical sentence
Those were the secrets$_i$ that [$_{RC}$ the politician attempted to cover up ___$_{pg}$] ultimately revealed ___$_i$ to the public.

iii. Gaps inside finite subject RC islands
a. Grammatical sentence
That was the famous dish$_i$ that [the chef [$_{RC}$ who had invented a special kind of spatula] won a lot of awards for ___$_i$.

23

b. Ungrammatical sentence
That was the famous dish$_i$ that [the chef [$_{RC}$ who had invented __$_i$]] won a lot of awards.

iv. Filler–gap dependencies into RC-complements
a. Grammatical sentence
She spoke a language$_i$ that I don't know that anybody else can speak __$_i$.

b. Ungrammatical sentence
She spoke a language$_i$ [that I don't know anybody else [$_{RC}$ that can speak __$_i$]]

The sentences constructed in this part of the experiment were modeled after the sentences in the SPR task. Different lexical items were used so the participants would not think that they read the same sentences again. The target items were grouped into four different categories as shown above. Category i) and iii) had four items each, whereas category ii) and iv) had two items each. There were twelve items in total, paired up with the same number of fillers.

3.2.3.2 Procedure

The judgment task was given after the self-paced reading task to prevent interference or priming from exposure to similar participants reading the ungrammatical similar sentences in the acceptability judgment task. In each trial, the participants were exposed to whole sentences and asked to judge them on a 7-point scale (1 = totally unacceptable, 7 = totally acceptable; only the endpoints were defined). Two example sentences to show the top and bottom ends of the scale were presented before the task began. The participants were instructed to judge each sentence based on whether they sounded like 'possible' sentences of English, while urging them to go with their initial instinct. No time limit was given, and the participants were told to take breaks if they were tired.

3.3 Ethical aspects

All participants were asked to participate voluntarily. Even though the experiment was hosted online, IP-addresses were not collected. Subjects were identified using a randomly assigned unique alpha-numeric code. As no identifying information was recorded, the experiment was considered exempt from reporting to the NSD (Norwegian Center for Research Data). Participants were told that they were able to withdraw from the experiment during their participation. No people below 18 years old participated in the experiment.

3.4 Predictions

In the following section, several possible outcomes of the collected data will be presented. First our predictions regarding active gap-filling will be presented. Then we will follow up with our predictions on island sensitivity in the L2 speakers.

3.4.1 Active vs. passive gap-filling

As previously noted, most research points to native speakers employing an active gap filling strategy when parsing sentences in English. Previous studies suggest that second language users also pursue an active gap filling strategy during online processing of filler-gap dependencies in English (Williams, Möbius & Kim, 2001; Omaki & Schulz, 2011; Kim, et al., 2015). Omaki & Schulz (2011) found that Spanish L1-English L2 speakers behaved similarly as their native control group, as the nonnative group also used an active gap filling strategy when parsing the non-island sentences (35). Similar items used in Omaki & Schulz have been replicated in this study. As nonnative speakers of English have been reported to follow an active gap filling strategy when parsing sentences in English, it is expected that our Norwegian speakers will also be active gap-fillers.

A plausibility mismatch paradigm is used to test whether the participants follow an active gap-filling strategy. The factor plausibility is evaluated in sentence tokens (a) implausible/non-island and (b) plausible/non-island in the experiment.

(35)
  a. /$_{Reg1}$ The city/ /$_{Reg2}$ that/ /$_{Reg\ 3}$ the author/ /$_{Reg\ 5}$ wrote/ /$_{Reg\ 6}$ regularly/ /$_{Reg\ 7}$ about/ /$_{Reg\ 8}$ was/ /$_{Reg\ 9}$ named/ /$_{Reg\ 10}$ after/ /$_{Reg\ 11}$ an explorer/

  b. /$_{Reg1}$ The book/ /$_{Reg2}$ that/ /$_{Reg\ 3}$ the author// $_{Reg\ 5}$ wrote/ /$_{Reg\ 6}$ regularly/ /$_{Reg\ 7}$ about/ /$_{Reg\ 8}$ was/ /$_{Reg\ 9}$ named/ /$_{Reg\ 10}$ after/ /$_{Reg\ 11}$ an explorer/

The '/' marks each region in the experimental sentences. Region 4 does not exist in the non-island conditions, as it is the relativizer 'who' that introduces the relative clause in the island condition. We expect to see a plausibility mismatch effect at the region 5, where the critical verb is situated. In this region, we expect that the participants try to assign the implausible filler (the city) as the object of the verb (wrote). This difference should be visible through an increased RT in the implausible constructions. The next region with the adverb (regularly) is also of interest as it serves as a site for possible spill-over effects after the critical verb region.

In the plausible condition (b), we expect no significant increase in RT in region 5 if the participants analyze *the book* as the object of *wrote*. It is, however, expected that the participants will need to reanalyse their syntactic structure later in the sentence when it becomes clear that the filler is actually the object of the preposition *about*. This should result in increased processing difficulty when the parser realizes that its initial structure was wrong. It is expected that this reanalysis effect will be shown through an increase in RT at or around region 7, which is where the actual gap site is located. This effect should not take place in the implausible constructions, as the parser has not committed to interpreting the filler as the object of the critical verb in region 5 because of implausibility.

The above expectations assume that the participants are active gap fillers. If no plausibility mismatch effects are present in any regions, we fail to reject the null hypothesis that participants are passive fillers.

3.4.2 Island Sensitivity

A comparison between the non-island and the island conditions in the SPR task will reveal whether the native Norwegian speakers posit a gap inside the RC island. If there is no significant difference in the size of the plausibility mismatch effect at or immediately after the first verb (*wrote*) between the island and non-island conditions in the self-paced reading task, this would suggest that active gap filling for the participants is not island sensitive. This might either be compatible with either the SSH or cross-linguistic transfer.

Kim et al. (2015) suggest that transfer from one's L1 might influence L2 processing, showing that native speakers of Korean (wh-in-situ) and Spanish (overt movement) parsed wh-dependencies in English slightly differently. Korean is classified as a wh-in-situ language, meaning that the wh-phrase does not move to form questions. On the other hand, Spanish is similar to English, as it also uses overt movement to form questions. As these languages have different syntactic properties in formation of filler-gap dependencies, it is possible that Norwegian's lack of RC island constraints can influence nonnative processing of islands as well. Based on these previous findings, it might be the case that the participants will not constrain their AGF strategy in the island conditions.

4 Results

This chapter presents the results from the acceptability judgment task and the self-paced reading experiment. The acceptability judgment ask was conducted to assess the participants' explicit (offline) knowledge of RC and subject island constraints. The self-paced reading experiment tested whether the participants were active gap fillers, and whether they were able to constrain their active gap filling strategy when processing a complex sentence containing an island.

4.1 Accuracy

Reading-time data from 6 participants whose accuracy on the comprehension questions in the self-paced reading task were < 75 % were excluded from the results. A low accuracy on the comprehension question could indicate that the participants did not understand enough of the sentences used in the experiment or were not completing the task carefully. In addition to participants with low accuracy scores, outliers were also excluded from the data. Outliers were categorized as any values under 200 ms and above 3000 ms (Keating & Jegerski, 2015, p. 20). Any reading times above 3000 ms might either reflect processing difficulty or distracted participants. The lower cutoff was determined on the basis that reading times below 200 ms are too fast for participants to have sufficiently processed the word, since lexical retrieval takes longer than 200ms.

      The six participants excluded in the self-paced reading task were also excluded from the judgment task.

4.2 Acceptability Judgment Task

The acceptability judgment task produced participants' numerical ratings ranging from 1 to 7. The mean ratings and standard errors for each condition will be presented.

4.2.1 Statistical analysis

Four paired t-tests were run in R (R Core Team, 2017) to test the difference between the mean rating of the paired grammatical and ungrammatical conditions (e.g., Finite-PG Ungrammatical vs. Finite-PG Grammatical). T-tests tell us whether the means of two groups are statistically significantly different from each other (Larsen & Marx, 2014, p. 458). If the means associated with two factor levels are equal, we fail to reject the null hypothesis (H_0: μ = μ_0 = 0). Our alternative hypothesis (H_1) is simply that the average rating in the two conditions tested against each other differed; the participants are aware of the grammatical differences between the unacceptable and acceptable constructions.

      A paired t-test was chosen as a suitable analysis tool as we are looking at each participant's mean rating in two different conditions. Each participant was measured twice, once for each sentence type. The participant's mean rating of every condition was calculated before the t-test was conducted. The ratings were the dependent variable, and the sentence types were the independent variable.

T- and p-values were interpreted during the statistical analysis. A high t-value ($|t| >$ 1) in combination with a low p-value ($p < .05$) would give us reason to reject the null hypothesis. The mean ratings for the filler items can provide a context in which we interpret the results. The grammatical fillers received a mean rating of 5.92 ($SD = 1.48$), and the ungrammatical fillers received a mean rating of 3.0 ($SD = 2.15$). The difference between these average ratings was significant ($t = -19.020$; $p < .001$).

4.2.2 Results

Mean acceptability ratings for each condition are presented in Table 4.1.

Table 4.1
Mean Acceptability Judgment Ratings and Standard Errors of the Mean

Condition	Grammatical	Ungrammatical
Filler	5.92 (0.06)	3.00 (0.09)
PG inside finite subject RC	4.51 (0.18)	2.88 (0.16)
PG inside infinitival subject RC	4.02 (0.27)	3.59 (0.26)
Gaps inside subject RCs	4.90 (0.16)	1.99 (0.12)
Gaps inside RC complements	4.98 (0.29)	3.82 (0.29)

Parasitic gaps inside finite subject relative clauses
The test revealed a significant difference between the grammatical and the ungrammatical condition ($t(50) = -6.791$, $p < .001$). The grammatical condition was rated higher than the ungrammatical condition. The effect size for this analysis ($|d| = 0.951$) was found to exceed Cohen's (1988) convention for a large effect ($d = .80$).

Gaps inside finite subject RC islands
The mean ratings of the grammatical condition were significantly higher than the ungrammatical condition ($t(50) = -11.469$, $p < .001$). The analysis revealed a very large effect size ($|d| = 1.606$).

Parasitic gaps inside infinitival subjects
In the infinite parasitic gap constructions, the mean ratings for the grammatical and ungrammatical condition were not significantly different ($t(50) = -1.138$, $p = .261$).

Filler-gap dependencies in RC-complements
In the complex sentences where a gap was situated inside an RC complement, the mean ratings in the grammatical condition were significantly higher than in the ungrammatical condition ($t(50)$ = -3.084, p = .003). The analysis revealed a small effect size ($|d|$ = 0.432).

As reported above, there was a significant effect between three out of four of the constructions. Although there was no significant effect in the sentences with a parasitic gap inside infinitival subjects, we did not use infinitival subject phrases in the test sentences in our self-paced reading task. This finding will be further discussed in Chapter 5.

It is important to note that offline results do not provide clear evidence of transfer and are not consistent with the SSH. This is because we are testing theories of real-time sentence processing, which does tap into explicit knowledge of language (as opposed to offline tasks). The results in the judgment task are used to check whether the participants are aware of the island constraints in English in an offline setting, which is vital for them being able to apply the constraint during sentence processing. The explicit knowledge of RC and subject islands is a prerequisite to test our main hypotheses.

4.2 Self-paced reading

4.2.1 Statistical analysis

The following hypotheses were formulated to test for active gap filling and island sensitivity in the participants.

Table 4.2

Hypotheses relevant for the SPR analysis

Hypothesis	Predicted results
H_0: The participants are not active gap fillers.	There is no effect of plausibility. RTs in the implausible and plausible sentences are not significantly different.
H_1: The participants are active gap fillers, but try to posit a gap inside non-islands and islands alike. This is consistent with the SSH and/or L1 transfer.	There is a main effect of plausibility, but no interaction effect. RTs in implausible and plausible sentences are significantly different.
H_2: The participants are active gap fillers, but will not try to posit a gap inside the island phrase.	There is an interaction effect between island and plausibility. RTs in the Implausible-NoIsland are significantly different from RTs in Plausible-Island, but there is no significant difference between RTs in the Implausible-Island and Plausible-Island conditions.

Statistical analysis used linear mixed effects models implemented using the lme4 (Bates, Maechler, Bolker & Walker, 2015) and lmerTest (Kuznetsova, Brockhoff and Christensen, 2017) packages in R (R Core Team, 2017).[15] Linear mixed models are similar to repeated measures two-way ANOVA. This type of linear model tests the differences between several means, with both fixed factors and random factors (Oehlert, 2010, p. 286). Fixed factors are effects preselected by the experimenter. Random factors are the opposite - effects not preselected by the experimenters (Larsen & Marx, 2014, p. 458).

Log-transformed reading times were the dependent variable. Log-transformation is used to make highly skewed distributions less skewed and closer to normally distributed, making the data consistent with the assumptions of the model that responses are normally distributed.

Models contained fixed effects of plausibility, island (Island v. no-Island), and their interaction. Models also included word-length as a fixed effect. Word length was a significant predictor in all models tested, but we do not report main effects of length because they are not relevant for our hypothesis. Models included random intercepts by subject and participants. No random-slopes were included because models with random slopes were either singular or did not converge.

4.2.1 Results

The mean overall accuracy for the comprehension questions was 86.6 % (*SD* = 3.4). The mean accuracy for the target sentences was 79.0 % (*SD* = 4.0) and the mean accuracy for the filler sentences was 90.7% (*SD* = 2.9). A high comprehension accuracy suggests that the participants were processing the sentences in a satisfactory way. Response accuracy is presented in Table 4.3.

Table 4.3

Accurate response table for the comprehension questions.

Condition	Mean	Standard Deviation	Standard error
Implausible-NoIsland	0.780	0.415	0.023
Plausible-NoIsland	0.848	0.359	0.020
Implausible-Island	0.723	0.448	0.024
Plausible-Island	0.810	0.393	0.021
Filler	0.907	0.290	0.006

---

[15] Whereas I conducted the t-tests on the results of the acceptability judgment task myself, my advisor Dave Kush implemented the linear mixed effects models. I was responsible for interpretation of all model results, however.

Figure 4.1

A plot of mean RT for each successive region (N=51). Islandhood is marked by color (blue for NoIsland and orange for Island conditions). The factor plausibility is marked by a straight line for the implausible conditions and a dotted line for the plausible condition.
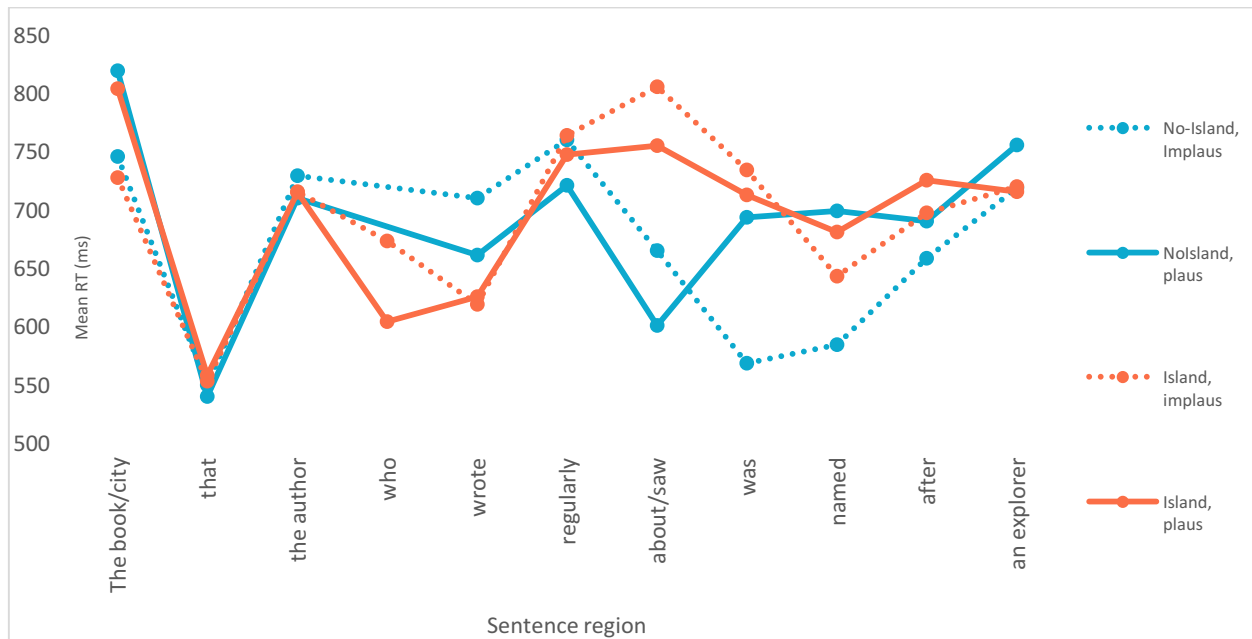
Average by-region SPR-reading times are plotted in Figure 4.1.

Up until the region with the critical verb *(wrote)*, there were no expected effects. At the critical verb there was a main effect of Island (t = -3.52, p < .001, -63.05ms): RTs in the NoIsland conditions were longer on average than in the Island conditions. There was not a significant effect of plausibility ($|t|$< 1) or a plausibility x Island interaction ($|t|$< 1).

At the post-verbal adverb ('regularly') there were no significant effects.

Two words after the critical verb ('about/saw') there was a marginally significant main effect of plausibility (t = 1.992, p = .047, -57.44ms). RTs were longer on average in Implausible conditions than in Plausible conditions. There was also a main effect of Island (t = -3.457, p < .001, 147.39ms): NoIsland conditions were read more quickly on average than Island conditions. Pairwise comparisons revealed that there was plausibility mismatch effect in the NoIsland conditions: The implausible-NoIsland condition was read more slowly than the Plausible-NoIsland condition (665ms vs. 600ms), (p = .0354). The trend towards a plausibility effect in the Island conditions was not significant. Thus, there was no reliable plausibility effect in the Island conditions.

Three words after the critical verb ('was') there were main effects of plausibility (t = -2.077; p =.0380, 52.24ms) and Island (t = -3.837, p < .001, 92.39ms). The main effect of Island indicates a spill-over effect from the previous region. There was also a significant plausibility x Island interaction (t = -2.719, p = .00664). This reflected that there were no

31

differences in RTs between the Plausible-Island and Implausible-Island conditions, but there was a large RT difference between the Plausible-NoIsland and Implausible-NoIsland conditions ($p < .001$). This reflected that the Plausible-NoIsland condition was read more slowly than the Implausible-NoIsland condition: the exact opposite direction of a typical plausibility mismatch effect. This interaction effect suggests that the participants did not expect a gap in the Plausible-NoIsland conditions and had to reanalyze the sentences. The same reanalysis effect was not present in the Implausible-NoIsland conditions as the participants were still anticipating a gap.

We interpret our results as supporting the hypothesis that the participants demonstrated evidence for active gap creation in L2. This allows us to reject our null hypothesis. Surprisingly, the results we predicted were delayed in our experiment. The implausible non-island condition was read more slowly than the plausible island condition at the filler integration region ('about/saw'), but there were no relevant effects for our hypothesis before this region. We interpret this as evidence for active gap filling as the plausibility mismatch effect seen in this region cannot be explained by other factors in this region. Therefore, the pairwise differences can be assumed to be related to previous regions. Further, the reversed plausibility mismatch effect in the region after ('was') also indicate AGF as the slowdown in the plausible conditions reveal that the participants had already linked the filler with the intermediate verb. Thus, they had to reanalyze their initial interpretation. The same effect was not displayed in the implausible conditions, as they were still waiting for a possible gap site.

The fact that there was a significant plausibility mismatch between the NoIsland conditions, but not in the Island conditions suggests that the L2 users did not pursue active gap creation inside the island. This allows us to reject H_1 as well. This pattern of results replicates the pattern found in the L1 Spanish-L2 English speakers in Omaki & Schulz's (2011) study.

Table 4.4
Mean RT for the critical regions.

Condition	wrote	regularly	about/saw	was
NoIsland, implausible	709.39	759.13	664.54	567.82
NoIsland, plausible	660.47	720.28	600.19	693.23
Island, implausble	618.56	762.99	805.13	733.38
Island, plausible	625.19	746.77	754.59	712.45

5 Discussion

In this chapter the results obtained in the acceptability judgment task and the self-paced reading experiment will be discussed. The chapter will discuss what implications the results will have for the theoretical background presented in chapter 2. First, we will discuss the findings. Second, we will discuss the SSH and L1 transfer. Lastly, we will discuss the weaknesses of this study and give suggestions for further research that needs to be conducted on the topic.

5.1 The aim of the study

The main focus of this study was to examine whether L2 users are able to constrain their active gap filling strategy when parsing island phrases. In particular, whether the L2 users would try to posit a gap inside an island phrase. Through a self-paced reading task, native Norwegian speakers' reading times of English filler-gap dependencies were recorded and compared. The filler-gap dependencies were manipulated through plausibility and islandhood (i.e. the island sentences had an additional RC island inside the subject phrase). Due to earlier findings which indicated that non-native speakers followed an active gap filling strategy (e.g. Williams, Möbius & Kim, 2001; Williams, 2006; Omaki & Schulz, 2011), it was expected that the participants in this study would exhibit the same behavior (i.e. a bias for creating a dependency at the earliest possible gap location).

We identified two reasons why Norwegian participants might try to actively fill gaps inside islands: First, the SSH predicts that L2 speakers should not display any sensitivity to island constraints during online sentence processing. This is because the SSH predicts that L2 users cannot build rich structural representations when parsing sentences in their target language, and they therefore have to rely more on semantic cues instead of syntactic cues. Second, results from Kim et al. (2015) suggested that island-insensitivity from L1 might transfer to L2: As Norwegian allows some island violations that are not allowed in English, we might have expected them to actively fill gaps inside constituents that are not islands in Norwegian.

5.2 Results

5.2.1 Active Gap-Filling

The results of the self-paced reading task show that Norwegians are active gap-fillers in L2 English. The reading time pattern revealed that the Norwegian speakers were sensitive to the plausibility manipulation at the earliest possible gap site, which indicated that they tried to actively predict where the gap site was when encountering the filler, and initially posited a gap at the intermediate verb. When compared with the native speaker data in Omaki & Schulz (2011, p. 578) we see that the processing pattern is similar, although the effects in this study occur later than expected. The strong reversed plausibility mismatch effect after the region containing the true gap site where the filler is actually integrated gave us additional (and arguably sufficient) evidence for active gap-filling in the Norwegian

speakers. These participants had to reanalyze their initial gap assignment when reaching the actual gap site in the sentence.

The results are consistent with earlier findings of non-native active filler-gap processing (Williams, Möbius & Kim, 2001; Omaki & Schulz, 2011; Kim et al., 2015). Together, these findings indicate that parsers attempt to form the shortest possible filler-gap dependency during online sentence processing, and that this strategy is used regardless of whether one processes the L1 or the L2.

5.2.2 Island sensitivity

5.2.2.1 Acceptability judgments

Participants were generally good at distinguishing grammatical filler-gap dependencies from ungrammatical filler-gap dependencies in English. Our results show that the speakers were aware of the RC island constraint in English and the fact that a parasitic gap inside a finite RC cannot be salvaged by a licit gap after the island phrase in English. Surprisingly, the participants could not differentiate the acceptable and unacceptable infinitival parasitic gap constructions. Although this finding is interesting, we believe that it should not affect the results as the experimental sentences in the SPR task only consisted of finite subject RC islands. The most relevant ratings were therefore the constructions involving parasitic gaps inside subject RCs, and gaps inside finite subject RCs. Those two constructions yielded the results with the biggest effect size, indicating that the participants had sufficient knowledge of the relevant island constraints.

5.2.2.2 Reading times

The results from the self-paced reading task gave evidence for the L2 users being sensitive to island constraints during real-time sentence processing. As discussed above, a plausibility mismatch effect was apparent in the non-island sentences. However, no plausibility mismatch effect was found in the island sentences, suggesting that the Norwegian speakers did not attempt to posit a gap inside the island domain, presumably by applying the island constraint to prevent the creation of an ungrammatical gap. These findings replicate the results from both the L1 control group and Spanish speakers in Omaki & Schulz (2011) and the Spanish participants in Kim et al. (2015). It further gives evidence of L2 speakers being able to construct advanced structural representations during online sentence processing.

5.3 Implications

5.3.1 The Shallow Structure Hypothesis

The fact that Norwegian speakers were able to block ungrammatical long-distance dependency formation in the self-paced reading experiment demonstrates that L2 speakers are able to build rich structural representations during online sentence processing. In addition, the L2 speakers are able to constrain their active search for a gap location when a 'possible', but ultimately ungrammatical gap is located inside an island. Thus, the results

cast doubt on the claim that advanced L2 speakers are unable to rely on syntactic cues during online sentence processing, which the SSH proposes. According to the SSH, the L2 speakers in this study should have relied on semantic cues (i.e. the semantic fit between the filler and the critical verb) over the syntactic island cue when trying to find a gap for the filler. Together with other studies on L2 processing the results here suggest advanced L2 speakers build syntactic structure in a similar way that native speakers do.

Omaki & Schulz (2011) suggest that the 'SSH could be maintained in a slightly weaker form (p. 584)' meaning that L2 speakers can be prone to construct shallow structures more often than native speakers do in some contexts. Specifically, when the L2 grammar is different from the L1 and transfer occurs. The results from this study provide some evidence against this claim (the 'island-less' Norwegians had no problem with the English subject RC-island constraint). However, the relative similarity between Norwegian and English in terms of word order (SVO/V2 vs SVO), typology, and overt movement when constructing RC and wh-questions might have helped Norwegians in constructing rich syntactic representations in L2. If L1 and L2 exhibited more basic syntactic differences, participants might perform less well. That could explain why the Korean speakers in Kim et al. (2015) were not sensitive to RC islands during online sentence processing.

5.3.2 L1 Transfer possibilities

Although Kim et al. (2015) found a possible transfer effect in their study, we did not find evidence that the Norwegian speakers in this study transferred 'island-insensitivity' from their L1 to L2. The Korean speakers in Kim et al. showed a plausibility mismatch effect in sentences containing a possible gap site inside an island, but the Spanish speakers in the same study did not. We could expect that native speakers of Norwegian, a language often referred to as an 'islandless language', would (in essence) be blind to the islands in this study. The results show that the Norwegian speakers respected the island constraints when parsing English island constraints, even though extraction out of RCs is allowed in Norwegian. In addition, even though a parasitic gap is allowed inside a finite subject phrase in Norwegian, the results indicated that the Norwegian speakers did not anticipate a gap after the island domain in the experimental sentences.

An important factor in explaining the difference between Korean and Norwegian behavior could proficiency and the age the participants started learning English. The participants in Kim et al. (2015) had a mean age of 11.1 (*SD* = 1.9, p. 391) of when they started learning English, but the participants in this study had a mean age of 6.3 (*SD* = 3.9). Although results from a recent study show that acquisition of syntactic properties can occur throughout childhood (Hartshorne et al., 2018), it may be harder to utilize grammar acquired after a certain age during online sentence processing. As discussed in Chapter 2.3.1, we depend more on the declarative memory with age (up until a certain point). Although the Korean speakers might have the explicit knowledge of RC islands stored in their declarative memory, they might not have proceduralized it which is needed for automatized actions (such as sentence processing). It could be argued that the Norwegian speakers were exposed to more English from a younger age, when they were more dependent on the procedural memory. Additionally, a lot of media consumed in Norway

(e.g. TV-series, films, video games, YouTube etc.) tend to be in English, even for children. The Norwegian speakers in this study also reported a high average consumption of English media. Similar data was not available from the Korean speakers, but it was reported that most of the participants had spent some time in an English-speaking country ($M = 3.6$ years, p. 391)[16].

5.4 Limitations and suggestions for further research

5.4.1 Experimental design

As we are comparing L2 processing to L1 processing, it would have been optimal to have a native speaker control group to compare our L2 speakers with. We reasoned that we did not need a control group since we could compare our results with the native speakers in Omaki & Schulz (2011) as that study involved sentences with similar construction. One small problem with such a comparison is that our test sentences did not have identical lexical items. Ideally, we would have used the same sentences in Omaki and Schulz, but the items were not accessible at the time of conduction. As we did not use identical sentences, it might have been better to have tested our own native speaker control group.

Another thing to note is that the expected plausibility mismatch effects were apparent after the regions we predicted to see the results in. We then raise the question of whether the spillover region (i.e. the adverbial region) should have been longer to differentiate between effects related to the critical verb inside the RC, and the region where the filler was retrieved. An example of a longer spillover region could be longer adverbs such as the ones used in Traxler & Pickering (1996, p. 465). It would be ideal to implement longer spill-over regions in any experiments with similar experimental design.

5.4.2 Participants' proficiency level

In the experiment, we attempted to determine the participants' proficiency level based on background information collected in the preliminary part of the experiment. All participants provided a high self-determined proficiency level and on average a high exposure/usage of English. However, we could have used a method (e.g. a C-test) to assess the participants' English skill more systematically. By being more stringent in identifying good and bad English speakers, we could run analyses on each group and check the differences, as proficiency is an important factor when testing L2 users.

5.4.3 Testing of transfer

We are merely basing our hypothesis of transfer on the basis that the participants find parasitic gaps inside subjects acceptable. For properties of one's L1 to transfer to L2, they actually need to be present within the parser. As we have previously stated, Engdahl (1983)

---

[16] Another reason for why the Norwegian speakers behaved differently might be due to differences in syntactic movement. Whereas Korean has mostly covert movement, and Norwegian has mostly overt movement. This is definitely something that should be researched further.

stated that individual differences have been reported in acceptability of island constraints in the Scandinavian languages (p. 8). It is not unreasonable to believe that there are individual differences among the fifty-one participants in this study. Additionally, as noted in chapter 2.1.3, there have been some cases where Norwegian speakers have rejected extraction from phrases that are believed to not elicit island effects (e.g. relative clauses in Taraldsen, 1982; Kush et al., 2018). Although it is widely believed that Norwegian speakers allow parasitic gaps within subject phrases, this should ideally have been assessed. We could have done this by administering a short acceptability judgment task on sentences with parasitic gaps inside subjects. By doing so, we would have made a stronger case for our hypothesis. Future research should follow up on this.

6 Conclusion and Summary

Second language processing may differ in many ways from native language processing, but the results from this study suggest that the processes that support filler-gap processing are qualitatively similar. Fifty-one native Norwegian speakers participated in a self-paced reading experiment to test their processing of English filler-gap dependencies. The results show that the Norwegian readers of English actively posit gaps when they are grammatical. However, the Norwegian speakers did not try to actively fill gaps inside subject RC islands when the result would be unacceptable in English. Additionally, the Norwegian speakers in the study were also tested in an offline acceptability judgment task to check whether they had knowledge of different English RC and subject islands constraints. The results show that the participants were aware of the constraints in an offline setting.

The possibility of L1 transfer was also tested in our experiment. As relative clauses are not considered islands in Norwegian, we assumed that this property might transfer to the Norwegian speakers' processing of English RCs. Norwegian also allows parasitic gaps inside subject RCs if there is a licit gap outside of the island domain. However, no indications of any L1 transfer effects were found in our results.

Even though the SSH posits that L2 users cannot build rich structural representations during language processing, the results from this experiment provide evidence against its claim; our nonnative speakers were processing the sentences similarly to the native speakers in Omaki & Schulz (2011). In sum, the results in this thesis provide evidence against the SSH, and to some extent, transfer effects. However, we do not completely reject the theory of L1 transfer. Instead, we propose that further research is needed on this topic as there are many factors (age, exposure, grammatical differences between the languages) that should be looked into to draw any conclusions on this matter.

7 References

Alexiadou, A., Law, P., Meinunger, A., & Wilder, C. (2000). *The Syntax of Relative Clause*s. Amserdam/Philadelphia: John Benjamins Publishing Company.

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

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Janua Linguarum 4.

Chomsky, N. (1977). *Essays on form and interpretation*. New York: North-Holland.

Chomsky, N. (1995). The Minimalist Program. Cambridge, MA: MIT Press.

Chomsky, N. (2000). Minimalist Inquiries: The Framework. In H. Lasnik, R. Martin, D. Michaels, & J. Uriagereka (Eds.), *Step by Step : Essays on Minimalist Syntax in Honor of Howard Lasnik (*pp. 89-156)*. Cambridge, MA: MIT Press.

Christensen, K. R., & Nyvad, A. M. (2014). On the nature of escapable relative islands. *Nordic Journal of Linguistics*, 37(1), 29-45. https://dx.doi.org/10.1017/S0332586514000055.

Clahsen, H., & Felser, C. (2006). Grammatical Processing in Language Learners. *Applied Psycholinguistics*, 27(1), 3-42. https://dx.doi.org/10.1017/S0142716406060024.

Clifton, C., & Frazier, L. (1989) Comprehending Sentences with Long-Distance Dependencies. In G. N. Carlson & M. K. Tanenhaus (eds.), *Linguistic Structure in Language* Processing (pp. 273-317). Dordrecht: Springer Netherlands.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, N.J.: Lawrence Erlbaum.

Crain, S., & Fodor, J. (1985). How can grammar help parsers? In D. Dowty, L. Kartunnen, & A. Zwicky (Eds.), *Natural Language Parsing.* Cambridge: Cambridge University Press.

Cunnings, I., Batterham, C., Felser, C. & Clahsen, H. (2010). Constraints on L2 learners' processing of wh-dependencies. In J. Jegerski & B. VanPatten (Eds.), *Research in Second Language Processing and Parsing* (pp. 87-109). Amsterdam: John Benjamins Pub. Co.

Cunnings, I. (2017). Parsing and Working Memory in Bilingual Sentence Processing. *Bilingualism: Language and Cognition*, 20(4), 659–678. https://dx.doi.org/10.1017/S1366728916000675.

Drummond, A. (2012). IbexFarm. Retrieved from http://spellout.net/ibexfarm/.

Eckes, T., & Grotjahn, R. (2006). A closer look at the construct validity of C-tests. *Language Testing*, 23(3), 290–325. https://dx.doi.org/10.1191/0265532206lt330oa.

Engdahl, E. (1983). Parasitic gaps. *Linguistics and Philosophy*, 6(1), 5-34. https://dx.doi.org/10.1007/BF00868088.

Fodor, J. D. (1979). Parsing strategies and constraints on transformations. *Linguistic Inquiry*, 9(3), 427-473.

Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition*, 68(1), 1-76. https://dx.doi.org/10.1016/S0010-0277(98)00034-1.

Gibson, E., & Warren, T. (2004). Reading-Time Evidence for Intermediate Linguistic Structure in Long-Distance Dependencies. *Syntax*, 7(1), 55-78. https://dx.doi.org/10.1111/j.1368-0005.2004.00065.x

Hopp, H. (2011). Individual differences in the second language processing of object-subject ambiguities. *Applied Psycholinguistics*, 36(2), 129-173. https://dx.doi.org/10.1017/S0142716413000180.

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

Juffs, A. (1998). Main Verb Versus Reduced Relative Clause Ambiguity Resolution in L2 Sentence Processing. *Language Learning*, 48(1), 107-147. https://dx.doi.org/10.1111/1467-9922.00034.

Kaiser, E. (2013). Experimental Paradigms in Psycholinguistics. In R. J. Podesva & D. Sharma's (Eds.), *Research Methods in Linguistics* (pp. 135-168). Cambridge: Cambridge University Press.

Kim, E., Baek, S., & Tremblay, A. (2015). The Role of Island Constraints in Second Language Sentence Processing. *Language Acquisition*, 22(4), 384-416. https://dx.doi.org/10.1080/10489223.2015.1028630.

Kush, D., Lohndal, T., & Sprouse, J. (2018). Investigating variation in island effects: A case study of Norwegian wh-extraction. *Natural Language & Linguistic Theory*, 36(3), 743-779. https://dx.doi.org/10.1007/s11049-017-9390-z.

Kush, D., Lohndal, T., & Sprouse, J. (2019) On the Island Sensitivity of Topicalization in Norwegian: An Experimental Investigation. *Language,* 95(3), 393-420. https://dx.doi.org/10.1353/lan.2019.0051

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

Larsen, R. J., & Marx, M. L. (2014). *An Introduction to Mathematical Statistics and Its Applications* (5th ed.). Harlow, England: Pearson Education Limited.

Mackey, A., & Gass, S. M. (2015). *Second Language Research: Methodology and Design*. Mahwah, N. J.: Lawrence Erlbaum.

Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in Second Language Sentence Processing. *Studies in Second Language Acquisition*, 27, 53-78. https://dx.doi.org/10.1017/S0272263105050035.

Maling, J., & Zaenen, A. (1982). A phrase-structure account of Scandinavian extraction phenomena. In P. Jacobson & G. K. Pullum's (Eds.), *The nature of syntactic representation* (pp. 229-282). Dordrecht: Springer Netherlands.

McKinnon, R., & Osterhout, L. (1996). Constraints on Movement Phenomena in Sentence Processing: Evidence from event-related potentials. *Language and Cognitive Processes*, 11(5), 495-524. https://dx.doi.org/10.1080/016909696387132.

Oehlert, G. W. (2000). *A First Course in Design and Analysis of Experiments*. University of Minnesota. Retrieved from: http://users.stat.umn.edu/~gary/book/fcdae.pdf.

Omaki, A., & Schulz, B. (2011). Filler-Gap Dependencies and Island Constraints in Second Language Sentence Processing. *Studies in Second Language Acquisition*, 33(4), 563-588. https://dx.doi.org/10.1017/S0272263111000313.

Phillips, C. (2006). The Real-Time Status of Island Phenomena. *Language,* 82 (4), 795-823.

Rizzi, L. (1982). *Issues in Italian Syntax*. Dordrecht: Foris.

Roberts, L., Gullberg, M., & Indefrey, P. (2008). Online pronoun resolution in L2 discourse: L1 influence and general learner effects. *Studies in Second Language Acquisition*,

30(3), 333–357. https://dx.doi.org/10.1017/S0272263108080480.

Ross, J. R. (1967). *Constraints on Variables in Syntax* (Doctoral dissertation). Massachusetts Institute of Technology (MIT).

Simensen, A. M. (2009). English in Scandinavia: a success story. In D. Wyse, R. Andrews and J. Hoffman (Eds.), *The Routledge International Handbook of English, Language and Literacy Teaching* (pp. 472-483). London: Routledge.

Sprouse, J. & Hornstein, N. (2013). Experimental syntax and island effects: Toward a comprehensive theory of islands. In J. Sprouse & N. Hornstein (Eds.), *Experimental syntax and island effects* (pp. 1-18). Cambridge: Cambridge University Press. https://dx.doi.org/10.1017/CBO9781139035309.001

Stowe, L. (1986). Evidence for on-line gap location. *Language and Cognitive Processes*, 1, 227–245.

Stowe, L. A. & Kaan, E. (2006). Developing an Experiment. Techniques and Design. Retrieved from: https://ufdcimages.uflib.ufl.edu/IR/00/00/55/17/00001/StoweKaan2006_s.pdf.

Paradis, M. (2009). *Declarative and Procedural Determinants of Second Languages*. Amsterdam: John Benjamin's Pub.

R Core Team (2017). R: A language and environment for statistical computing. Vienna, Austria.: R Foundation for Statistical Computing. Retrieved from https://www.Rproject.org/.

Taraldsen, K. T. (1982). Extraction from Relative Clauses in Norwegian. In E. Engdahl, & E. Ejerhed (Eds.), *Readings on Unbounded Dependencies in Scandinavian Languages*. Stockholm: Almqvist & Wiksell International.

The Norwegian Directorate of Education and Training. (1997). Læreplanverket for den 10-årige grunnskolen. Oslo: Nasjonalt læremiddelsenter. Retrieved from: https://www.nb.no/nbsok/nb/f4ce6bf9eadeb389172d939275c038bb?lang=no#231.

The Norwegian Directorate of Education and Training. (2013). English Subject Curriculum. Retrieved from: http://data.udir.no/kl06/rest_/ENG1-03.pdf?lang=eng.

Torrego, E. (1984). On inversion in Spanish and some of its effects. *Linguistic Inquiry*, 15(1), 103–129.

Traxler, M. J. & Pickering, M. J. (1996). Plausibility and the Processing of Unbounded Dependencies: An Eye-Tracking Study. Journal of Memory and Language, 35(3), 454-475. https://dx.doi.org/10.1006/jmla.1996.0025.

Traxler, M. J. (2012). Introduction to Psycholinguistics. Understanding Language Science. West Sussex: Wiley-Blackwell.

Ullman, M.T. (2001). The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language & Cognition*, 4(2), 105-122. http://dx.doi.org/10.1017/S1366728901000220.

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition. The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory and practice* (pp. 151-178). Washington, D. C.: Georgetown University Press.

Williams, J. N. (2006). Incremental interpretation in second language sentence processing. *Bilingualism: Language and Cognition*, 9, 71-88. https://dx.doi.org/10.1017/S1366728905002385.

Williams, J. N., Möbius, P. & Kim, C. (2001). Native and non-native processing of English wh-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics,* 22(4), 509–540. https://dx.doi.org/10.1017/S0142716401004027.

Witzel, J., Witzel, N. & Nicol, J. (2012). Deeper than shallow: Evidence for structure-based parsing biases in second-language sentence processing. *Applied Psycholinguistics*, 33(2), 419-456. https://dx.doi.org/10.1017/S0142716411000427.

Appendices

Appendix 1 – Relevance of the study for the teaching profession

Appendix 2 - Questionnaire

Appendix 3 – Materials for the self-paced reading task.

Appendix 4 – Materials for the acceptability judgment task.

Appendix 5 – Additional background information on the participants.

Appendix 1 – Relevance of the Study for the Teaching Profession

This thesis has examined how L2 users of English process complex sentences. In particular, which strategies advanced L2 speakers utilize when parsing filler-gap dependencies. As a future English teacher, it can be helpful to have an idea of the strategies L2 users incorporate when processing language. Understanding that highly proficient L2 users are essentially native-like in their processing of English sentences could be useful in language teaching.

Likewise, studying topics of L1 transfer and properties of the Norwegian language has given me insight of difficulties Norwegian speakers might experience when working with English as a second language. Although transfer might not be as apparent in advanced sentence processing (as our results indicate) it is helpful to be aware of certain grammatical properties of Norwegian which may or may not transfer, regardless of proficiency level. That way, it is easier to recognize common mistakes Norwegian speakers (and in general L2 users) make when using English.

The process of writing this thesis was a valuable experience in itself. As a future language teacher, I will help students with 'written communication', as listed in The English Subject Curriculum. The long process of writing and re-writing will help me guide future students during their own writing projects, and the process of receiving feedback from my advisor will surely help me give valuable feedback to my own students in the future. This is helpful for the aim 'write different types of texts with structure and coherence suited to the purpose and situation' for upper secondary school, *VG1* (The English Subject Curriculum, 2013).

Further, I had the chance to work with different softwares. I was able to work in-depth with both Microsoft Word and Google Docs. as we utilized different tasks I was unfamiliar with, such as the functions under the Review-button and how to simply format a text. I also had the chance to participate in a LingPhil course on statistics, which gave me valuable information about the statistics program R. Although I'm not sure I will get to use all the new features I've learned in my future as a teacher, I believe that simply getting more acquainted with softwares will help me develop my own digital skills, which are becoming increasingly important today in the classroom.

Additionally, working with statistics has given me a foundation for incorporating math in my English classes. As 'tverrfaglighet', or interdisciplinary cooperation across different subjects is becoming increasingly important after the new curriculum is implemented this fall (*Fagfornyelsen*), knowledge in other subjects in school is important for incorporating different subjects in my English classes. Moreover, numeracy is one of the basic skills (*grunnleggende ferdigheter*) which is supposed to be incorporated in each subject in school. Thus, I believe that working with statistics will surely be of use in my future as an English teacher.

Appendix 2 – Questionnaire

Hi!

Thank you for participating in my thesis experiment.
Before we start we need to know a bit more about you and your language
background. Please answer the following questions:

Age	
Gender	
Is Norwegian your native language?	
Where are you from in Norway?	
Is Norwegian the native language(s) of your parents?	
Did you speak any other languages as a baby or young child?	◯ No ◯ Yes
When did you start learning English?	◯ 4 years old or younger ◯ 5-7 years of age ◯ 8-10 years of age ◯ 11-15 years of age ◯ Over 15 years of age
How would you rate your proficiency in English on a scale of 1-7(1 = very low, 7 = native fluency)?	1 ◯ ◯ ◯ ◯ ◯ ◯ ◯ 7
How many hours on average a week do you use English? (school assignments, chat, conversation, etc.)	◯ Not much (0-1 hrs a week) ◯ Some (1-3 hours a week) ◯ 3-5 hours ◯ 5-10 hours ◯ 10-20 hours ◯ 20+ hours
How much spoken English media are you exposed to on a regular day? (TV-series/livestreams/videos, Internet pages or apps in English, books written in English, lectures in English etc…)	◯ Less than 1 ◯ 1-2 hours ◯ 3-5 hours ◯ 5+ hours
Do you use English regularly with any close friends or family?	◯ No ◯ Yes
Have you ever lived for more than month in an English-speaking country?	◯ No ◯ Yes
Have you completed any coursework in English at University? If so, how many study points?	◯ No coursework in English ◯ Less than 20 study points ◯ 20-40 study points ◯ 40-60 study points ◯ 60+ study points

Appendix 3 – Self-Paced Reading Items

A: implausible, nonisland; B: plausible, nonisland; C: implausible, island; D: plausible, island.

#		Sentence	Comp. Q.
1	A	The city that the author wrote regularly about was named after an explorer.	Was the author named after an explorer?
1	B	The book that the author wrote regularly about was named after an explorer.	
1	C	The city that the author who wrote regularly saw was named after an explorer.	
1	D	The book that the author who wrote regularly saw was named after an explorer.	
2	A	The hospital that the patient read slowly about made her feel better.	Did the hospital/novel make her feel better?
2	B	The novel that the patient read slowly about made her feel better.	
2	C	The hospital that the patient who read slowly liked made her feel better.	
2	D	The novel that the patient who read slowly liked made her feel better.	
3	A	The girl that the lady knitted carefully with was right next to the TV.	Was the girl/yarn right next to the TV?
3	B	The yarn that the lady knitted carefully with was right next to the TV.	
3	C	The girl that the lady who knitted carefully patted was right next to the TV.	
3	D	The yarn that the lady who knitted carefully patted was right next to the TV.	
4	A	The monument that the tourists read eagerly about was in Spain.	Did the tourists look at a map?
4	B	The brochure that the tourists read eagerly about was in Spain.	
4	C	The monument that the tourists who read eagerly looked at was in Spain.	
4	D	The brochure that the tourists who read eagerly looked at was in Spain.	
5	A	The sign that the motorist drove quickly past was on the left side of the road.	Was the motorist on
5	B	The car that the motorist drove quickly past was on the left side of the road.	

5	C	The sign that the motorist who drove quickly saw was on the left side of the road.	the left side of the road?
5	D	The car that the motorist who drove quickly saw was on the left side of the road.	
6	A	The spatula that the chef cooked regularly with was bought at Walmart.	Was/were the spatula/ingredients bought at Walmart?
6	B	The ingredients that the chef cooked regularly with were bought at Walmart.	
6	C	The spatula that the chef who cooked regularly used was bought at Walmart.	
6	D	The ingredients that the chef who cooked regularly used were bought at Walmart.	
7	A	The tribe that the linguist spoke regularly about is currently endangered.	Did the linguists speak about/research a tribe/dialect?
7	B	The dialect that the linguist spoke regularly about is currently endangered.	
7	C	The tribe that the linguist who spoke regularly researched is currently endangered.	
7	D	The dialect that the linguist who spoke regularly researched is currently endangered.	
8	A	The editor that the journalist wrote frequently for was featured in a local magazine.	Was the journalist featured in the magazine?
8	B	The advice column that the journalist wrote frequently for was featured in a local magazine.	
8	C	The editor that the journalist who wrote frequently liked was featured in a local magazine.	
8	D	The advice column that the journalist who wrote frequently liked was featured in a local magazine.	
9	A	The cowboy that the farmer rode eagerly with was shot by a vigilante.	
9	B	The horse that the farmer rode eagerly with was shot by a vigilante.	
9	C	The cowboy that the farmer who rode eagerly cared for was shot by a vigilante.	Was the farmer shot?
9	D	The horse that the farmer who rode eagerly cared for was shot by a vigilante.	

10	A	The books that the singer sang beautifully from were well-known.	Was the singer well-known?
10	B	The opera that the singer sang beautifully from was well-known.	
10	C	The books that the singer who sang beautifully heard were well-known.	
10	D	The opera that the singer who sang beautifully heard was well-known.	
11	A	The town that the pilot flew gracefully over was full of people.	Were there a lot of people in/on the town/plane?
11	B	The plane that the pilot flew gracefully over was full of people.	
11	C	The town that the pilot who flew gracefully liked was full of people.	
11	D	The plane that the pilot who flew gracefully liked was full of people.	
12	A	The family that the maid cleaned regularly for was under a lot of pressure.	Was the maid under a lot of pressure?
12	B	The hotel that the maid cleaned regularly for was under a lot of pressure.	
12	C	The family that the maid who cleaned regularly hated was under a lot of pressure.	
12	D	The hotel that the maid who cleaned regularly hated was under a lot of pressure.	
13	A	The pot that the woman cooked skillfully with was of high quality.	Was/were the pot/beans of high quality?
13	B	The beans that the woman cooked skillfully with were of high quality.	
13	C	The pot that the woman who cooked skillfully bought was of high quality.	
13	D	The beans that the woman who cooked skillfully bought were of high quality.	
14	A	The friend that the man hunted stealthily for was running away.	Did the man run away?
14	B	The deer that the man hunted stealthily for was running away.	
14	C	The friend that the man who hunted stealthily saw was running away.	
14	D	The deer that the man who hunted stealthily saw was running away.	
15	A	The country that the spy killed mercilessly for gave him a lot of money.	

15	B	The general that the spy killed mercilessly for gave him a lot of money.	Was the spy given any money?
15	C	The country that the spy who killed mercilessly betrayed gave him a lot of money.	
15	D	The general that the spy who killed mercilessly betrayed gave him a lot of money.	
16	A	The shampoo that the hairdresser read extensively about was thrown away later that day.	Did the hairdresser read about/buy a shampoo/magazine?
16	B	The magazine that the hairdresser read extensively about was thrown away later that day.	
16	C	The shampoo that the hairdresser who read extensively bought was thrown away later that day.	
16	D	The magazine that the hairdresser who read extensively bought was thrown away later that day.	
17	A	The blocks that the boy played happily with were incredibly expensive.	Were/Was the blocks/guitar incredibly cheap?
17	B	The guitar that the boy played happily with was incredibly expensive.	
17	C	The blocks that the boy who played happily broke were incredibly expensive.	
17	D	The guitar that the boy who played happily broke was incredibly expensive.	
18	A	The actor that the editor wrote a lot about was nominated for an award.	Was the editor nominated for an award?
18	B	The script that the editor wrote a lot about was nominated for an award.	
18	C	The actor that the editor who wrote a lot liked was nominated for an award.	
18	D	The script that the editor who wrote a lot liked was nominated for an award.	
19	A	The plate that the man ate greedily from cost him a lot of money.	Was/were the plate/appetizers expensive?
19	B	The appetizers that the man ate greedily from cost him a lot of money.	
19	C	The plate that the man who ate greedily bought cost him a lot of money.	
19	D	The appetizers that the man who ate greedily bought cost him a lot of money.	

20	A	The students that the employee cooked enthusiastically for were prepared for a presentation.	Was the employee prepped for a presentation?
20	B	The meals that the employee cooked enthusiastically for was prepared for a presentation.	
20	C	The students that the employee who cooked enthusiastically saw were prepared for a presentation.	
20	D	The meals that the employee who cooked enthusiastically saw was prepared for a presentation.	
21	A	The student that the musician practiced a lot with could be found in the school's music department.	Did the musician practice a lot?
21	B	The instruments that the musician practiced a lot with could be found in the school's music department.	
21	C	The student that the musician who practiced a lot admired could be found in the school's music department.	
21	D	The instruments that the musician who practiced a lot admired could be found in the school's music department.	
22	A	The woman that the critic spoke enthusiastically about was very interesting to listen to.	Was the critic interesting?
22	B	The lyrics that the critic spoke enthusiastically about was very interesting to listen to.	
22	C	The woman that the critic who spoke enthusiastically knew was very interesting to listen to.	
22	D	The lyrics that the critic who spoke enthusiastically knew was very interesting to listen to.	
23	A	The house that the dog dug excitedly in was his favorite hiding spot.	Did the dog dig/see a house?
23	B	The tunnel that the dog dug excitedly in was his favorite hiding spot.	
23	C	The house that the dog who dug excitedly saw was his favorite hiding spot.	
23	D	The tunnel that the dog who dug excitedly saw was his favorite hiding spot.	
24	A	The machine that the body builder trained passionately with will run until it's forced to stop.	Will the body builder run until he is
24	B	The athlete that the body builder trained passionately with will run until he's forced to stop.	

24	C	The machine that the body builder who trained passionately loved will run until it's forced to stop.	forced to stop?
24	D	The athlete that the body builder who trained passionately loved will run until he's forced to stop.	

Appendix 4 – Judgment Task Items

Parasitic gaps inside finite subject RC islands

#	Grammatical	Ungrammatical
1	Those were the animals that the zookeeper that looked after the amphibians had cared for __ before he lost his job	Those were the animals that the zookeeper that looked after __ had cared for __ before he lost his job.
2	That was the school that the man that tried to burn down his house had attended __ as a kid.	That was the school that the man tried to burn down __ had attended __ as a kid.
3	That was the university that the woman who had donated money to a charity had studied at __ for her law degree.	That was the university that the woman who had donated money to __ had studied at __ for her law degree.
4	That was the murder case that the law students who learned about the constitution had discussed __ in preparation for their exam.	That was the murder case that the law students who learned about __ had discussed __ in preparation for their exam.

Parasitic gaps inside infinitival subject RC islands

#	Grammatical	Ungrammatical
5	Those were the secrets that the attempts to cover up __ultimately revealed __ to the public.	Those were the secrets that the politician attempted to cover up __ ultimately revealed __ to the public.
6	Those were the endangered species that the attempts to protect __ actually killed off __.	Those were the endangered species that the ranger attempted to protect __actually killed off __.

Gaps inside subject RC islands

#	Grammatical	Ungrammatical
7	That was the famous dish that the chef that had invented a special kind of spatula won a lot of awards for __.	That was the famous dish that the chef that had invented __ won a lot of awards.
8	That was the plant that the gardener that had watered the orchids placed __ out on the patio.	That was the plant that the gardener that had watered __ placed some other pots out on the patio.
9	That was the TV that the man that had bought a new remote loved to watch the news on __.	That was the TV that the man that had bought __ loved to watch the news.
10	That was the book that the scholar that had read the Odyssey wrote his thesis on __.	That was the book that the scholar that had read __ wrote his thesis.

Gaps inside RC complements

#	Grammatical	Ungrammatical
11	She spoke a language that I don't know that anybody else can speak __.	She spoke a language that I don't know anybody else that can speak __.
12	He recommended a book that I know a lot of people have read __.	He recommended a book that I know a lot of people who have read __.

Appendix 5 – Additional demographic/background info on the participants

Yes/no questions on participants' linguistic background.

Answers (in %)	Did you speak any other languages as a baby or young child?	Is Norwegian the native language(s) of your parents?	Do you use English regularly with any close friends or family?	Have you ever lived for more than a month in an English-speaking country?
Yes	9.80%	98.04%	37.25%	39.22%