Centre for Languages and Literature
Master's Thesis, 30 hp (SPVR01)
English Language and Linguistics
August 2012

Processing filler-gap dependencies in an L2:

An Event-Related Potential study

Andrea Schremm

Supervisors:
Carita Paradis
Mikael Roll

Acknowledgments

I would like to thank my supervisors, Carita Paradis and Mikael Roll, for their invaluable feedback and guidance during the research process and the writing of this thesis.
I am grateful to Damon Tutunjian for being the voice of my stimuli and to Sabine Goßelke for helping me finding participants. I also wish to thank my neighbours, who were not here during this summer and Patrik, who always was.

Abstract

The present study investigated second language (L2) learners' processing of filler-gap dependencies and tested the implications of Clahsen and Felser's (2006) *shallow structure hypothesis* according to which L2 learners underuse syntactic information during the processing of these sentence types. Advanced L2 learners of English (native Swedish speakers) listened to English sentences with object-relative, subject-relative and finite complement clauses. Event-related potentials were recorded at the word *that*, the embedded verb and the following word: a prepositional phrase, a semantically congruent or incongruent noun phrase or an extra noun phrase ungrammatically occupying the canonical object position after the verb in object-relative clauses. Compared to correct sentences, the extra noun phrase elicited a late positivity with a central-posterior distribution (P600). No increased negativity indicating problems with semantic processing (N400) was observed, suggesting that the extra noun phrase was initially interpreted as the object of the verb, and the integration of the clause-initial noun phrase (the *filler*) with the verb was delayed. The syntactic violation involved in this interpretation led to syntactic repair processes reflected in the P600 effect. The relative pronoun *that* in object-relative clauses gave rise to an enhanced negativity over anterior and central sites probably indexing storage of the filler in working memory. Finally, semantically anomalous nouns yielded a P600 but no N400 effect in grammatically correct complement clauses, which was interpreted to reflect the conflict between the relations dictated by the syntactic structure and the alternative thematic relationships that might be established to make sense of these sentences. The observation that participants relied on both syntactic and semantic information during sentence comprehension is not consistent with Clahsen and Felser's proposal. The results of this study suggest that learners might differ from native speakers in terms of processing speed, which potentially affects the relative order and nature of interaction of different subprocesses involved in sentence comprehension.

Contents

List of figures

1. Introduction

To provide a comprehensive account of learners' competence of a second language (L2 for short), we need to explain how learners use their linguistic knowledge during real-time sentence comprehension. Studies employing measures that provide information on the way sentence processing unfolds in time have revealed differences between L2 learners and native speakers (NSs for short) also in cases where they performed similarly on off-line language tasks (e.g.: Felser, Cunnings, Batterham, & Clahsen, 2012; Marinis, Roberts, Felser, & Clahsen, 2005; Papadopoulou & Clahsen, 2003; Roberts & Felser, 2011). Based on the results of these studies, Clahsen and Felser (2006) proposed an account of sentence processing, the *shallow structure hypothesis* (SSH for short), which claims that L2 sentence comprehension differs from native language processing in that the former relies on shallow parsing. According to Clahsen and Felser's definition, shallow parsing involves the segmentation of the linguistic input into meaningful units and the establishment of semantic relationships between the chunks primarily based on semantic roles, probabilistic information, pragmatic fit and world knowledge. The authors argue that NSs, as opposed to L2 learners, quickly integrate bottom-up and top-down information during parsing: they construct a detailed syntactic representation of the incoming word string while they also make use of a wide-range of semantic-pragmatic information sources. L2 learners are thus assumed to be less sensitive to syntactic cues than NSs, which they compensate for by relying more on other types of information (Clahsen & Felser, 2006; Roberts & Felser, 2011). While Clahsen and Felser do not make their assumptions explicit, their treatment of the topic suggests that syntactic information in this context is understood as cues to interpretation provided by the computation of a hierarchical phrase structure, whereas the information sources learners are assumed to make use of involve world knowledge and other types of information extracted from the input without constructing a phrase structure representation. The authors, therefore, assume that NSs create mental representations of sentences by organizing words into hierarchical units. From this perspective, the possible combinations of word categories are constrained by the grammar of the language, and word category is defined as a group of words that pattern together on the basis of their morphosyntactic behaviour regardless of their semantic properties. It is important to mention that not all accounts of native sentence processing presuppose the computation of phrase structure representations, and many grammatical approaches do not even use constructs such as phrase structure or word category in their description of linguistics phenomena. Nevertheless, since

the above-described assumptions and terminology are widespread in the context of language processing research (Frisch, Hahne, & Friederici, 2004) they are adopted in this paper as well due to practical considerations.

Commentaries on Clahsen and Felser's (2006) article suggest that the SSH is a much welcomed contribution to the field of second language acquisition and psycholinguistics as it draws attention to an important new line of investigation. However, most commentators emphasized the need for further research to test and specify the claims of the SSH. One linguistic phenomenon that provides an appropriate testing ground for the hypothesis involves the so-called *filler-gap dependencies*; Clahsen and Felser (2006) formulated specific predictions concerning the way learners process these sentence types. Filler-gap dependencies create a "separation between the position where a phrase is pronounced and the verb (or other head) which determines its thematic role" (Phillips & Wagers, 2007, p. 743). Sentences involving filler-gap dependencies in English include relative clauses such as (1):

(1) The receptionist that the painter scared by accident answered the phone.

The noun phrase (*the receptionist*) is the so-called *filler*, which appears sentence-initially in a position that is different from its canonical location in the sentence (the so-called *gap*) after the verb *scared*. In order to account for the real-time comprehension of these structures, it is necessary to identify the processes involved in establishing a relationship between the filler and the element that determines its thematic role (the verb *scared* in (1)). Certain accounts of filler-gap processing claim that NSs construct a structurally present gap in the canonical position of the filler, and integration with the verb is mediated through linking the filler to this gap (e.g. Nicol & Swinney, 1989; see further in Section 4). By contrast, Clahsen and Felser (2006) proposed that L2 learners establish a direct semantic association between the filler and the verb.

This thesis examines L2 learners' processing of sentences with filler-gap dependencies such as (1) with the aim of testing the implications of the SSH and investigating the real-time sentence comprehension mechanisms employed by L2 learners of English. The main research question concerns the issue of whether L2 learners underuse syntactic information during the on-line processing of filler-gap dependencies and establish a direct semantic association between the filler and its subcategorizer as predicted by the SSH. This question is addressed with the help of the *event-related brain potential* (ERP) technique as its high temporal resolution provides an accurate measurement of the relative timing of

different processing activities during language comprehension. The study focuses on the ERP responses that L2 learners of English (Swedish native speakers) display while they listen to English relative clauses including sentences in which the canonical position of the clause-initial object is ungrammatically occupied by an additional noun phrase. The findings are evaluated against results obtained with native English speakers in a previous study using a similar design (Hestvik, Maxfield, Schwartz, & Shafer, 2007). In order to gain information about the working memory and integration processes involved in linking the filler to the verb, it is also examined how learners process the word that marks the onset of a relative clause (*that*) and the embedded verb at which integration with the filler is assumed to take place. The obtained ERP effects are compared to responses that NSs have been reported to display.

The thesis is organized as follows. In Section 2, further information on Clahsen and Felser's (2006) shallow structure hypothesis is presented along with a brief summary of the experimental findings that provided the empirical basis for the SSH. Section 3 introduces the relevant ERP components observed with native and L2 language comprehension followed by an overview of previous research on the processing of filler-gap dependencies by NSs and L2 learners. The study that has been conducted as part of this master's thesis is presented in the subsequent sections. After a discussion of the obtained results, the most important findings and conclusions are summarized.

2. The shallow structure hypothesis

Clahsen and Felser (2006) formulated their proposal of the SSH on the basis of a series of studies that investigated the way L2 learners and NSs process structurally ambiguous sentences and syntactic dependencies. For example, Felser, Marinis and Clahsen (2003) and Papadopoulou and Clahsen (2003) observed that advanced L2 learners of English and Greek with a variety of first language backgrounds differed from NSs in their processing of relative clause attachment ambiguities. The experimental sentences involved a temporary ambiguity related to the fact that the relative clause could be potentially associated with either an earlier noun phrase (NP for short) or a later one. For example in (2), the relative clause *who was reading the letter* can be taken to modify either *the secretary* or *the professor.*

(2) The dean liked the secretary of the professor who was reading the letter. (Felser et al., 2003, p. 480)

L2 learners performing a self-paced reading task were not consistent in their associations of the relative clause with either the earlier or the later NP indicating that they did not use any of the phrase structure-based parsing principles such as *predicate proximity* and *recency preference*, which have been proposed to guide the resolution of attachment ambiguities in NSs (e.g. Gibson, Pearlmutter, Canseco-Gonzalez, & Hickock, 1996). Predicate proximity entails a preference for attaching ambiguous modifiers structurally as close to a predicate phrase as possible, and recency preference favours attachment to the phrase that has been more recently processed. The learners, nevertheless, displayed sensitivity to the semantic properties of the preposition that linked the two NPs. Similarly to NSs, they generally preferred disambiguation with the second NP in case of a thematic preposition such as *with*. The apparent absence of structure-based locality principles in L2 processing was taken to support the assumption that learners fail to construct syntactic representations with enough detail to underlie the operation of structure-based parsing strategies (Clahsen & Felser, 2006).

In a subsequent study, Roberts and Felser (2011) obtained findings that were interpreted to reflect L2 learners' greater reliance on plausibility information. The researchers investigated the processing of garden-path sentences that were temporarily ambiguous concerning the role of an NP as the object of the preceding verb or the subject of a following clause: in sentence (3), *the boss* is a plausible object of the verb *warned* while *crimes* is implausible. The manipulation of the plausibility of the NP as the object of the verb had a greater effect on the reading times of learners as compared to NSs. Learners took longer to read NPs that were implausible objects than those that were plausible objects. After the disambiguating verb (*destroy*), which made it clear that the NP was not the object of the preceding verb, reading times were longer for sentences containing plausible object NPs suggesting difficulties with reanalysing the initial interpretation.

(3) The inspector warned the boss (crimes) would destroy very many lives. (p. 304)

Clahsen and Felser (2006) also discuss the possibility that factors other than parsing strategies might also be accountable for the differences found in real-time native and L2 language processing, and they draw the conclusion that the findings of the reviewed studies are most consistent with the SSH. They argue that incomplete grammatical knowledge cannot explain their findings as the learners in Felser et al.'s (2003) and Papadopoulou and Clahsen's (2003) studies performed in the range of NSs on off-line tasks. In addition, they claim that the observed results were unlikely to depend solely on differences in native and L2

processing speed and working memory resources. They found that children acquiring their first language (L1 for short) tend to prioritize syntactic information to deal with processing difficulties resulting from their limited working memory capacity. From this perspective, L2 learners would be expected to rely on the same strategy to compensate for the shortage of working memory resources. Moreover, learners were found to differ from both high and low working memory span NSs in their processing of indirect object dependencies, and the pattern of results obtained with them were not modulated by differences in working memory capacity (Felser & Roberts, 2007; see further in Section 4). Finally, they argue that L2 learners do not seem to transfer L1 parsing strategies to L2 comprehension as learners with different L1 backgrounds displayed highly similar non-native processing patterns irrespective of whether their L1 was characterized by a similar or different preference concerning the investigated phenomena (Felser et al., 2003; Marinis et al., 2005; Papadopoulou & Clahsen, 2003).

With regard to the reasons why learners would be restricted to shallow parsing in the L2 if they are assumed to be able to compute full representations in their L1, Felser and Clahsen (2009) mention Ullman's (2001) declarative/procedural model as a possible explanation. Ullman's proposal distinguishes between a procedural and a declarative memory system. The declarative system underlies knowledge about facts and events, and the information represented in this way is at least partly explicit. It is assumed that the lexicon, which includes mainly arbitrary knowledge about words in a language, is also subserved by declarative memory. The procedural memory system is implicit and it supports the acquisition and control of sequences, skills and habits involving the computation of grammatical structures. The two systems interact and compete in the learning and processing of information: enhanced learning in one system might suppress the function of the other. Ullman argues that as a result of certain neurological maturational changes, late L2 learners show a greater reliance on the declarative system than NSs, even for syntactic operations that depend on procedural memory in NSs. As oestrogen is assumed to enhance the functioning of declarative memory, increase in the level of this hormone in both genders during childhood and adolescence might be partly responsible for these changes. In case of late L2 learners, compositional forms such as phrases in the L2 are stored in declarative memory, and, therefore, structures that are not easily memorized in terms of chunks and stored schemas, such as long-distance dependencies, might cause problems. Nevertheless, practice is expected to lead to an increasingly native-like reliance on procedural memory (Ullman, 2006). In their original proposal of the SSH, Clahsen and Felser (2006) found this last point problematic as

they argued that even highly proficient L2 learners seemed to differ from NSs in their processing strategies. Felser and Clahsen (2009) still maintain that high proficiency does not lead to native-like sentence processing mechanisms in an L2; it is not clear how the researchers consolidate these claims with the predictions of Ullman's theory that experienced L2 speakers depend on procedural memory for grammatical processing in the same way as NSs do.

While research conducted by Clashen, Felser and their colleagues obtained results consistent with the SSH, there are also a number of studies reporting findings that question some of the basic assumptions of the SSH. For example, Frenck-Mestre and Pynte (1997) observed L1 influence on L2 syntactic processing: it took longer for the learners to process verbs that were transitive in the L2 and intransitive in their native language. At the same time, the authors concluded that learners seem to perform a syntactic parse of the sentence in the same way as natives do. Similarly, Hopp (2006) reported that the highly proficient L2 learners of their study did not appear to differ from the NSs in their processing routines and displayed native-like use of syntactic features during structural reanalysis. A further issue concerns the fact that the SSH has so far been tested only with a very restricted range of sentence types, and the assumption that learners are generally confined to shallow parsing principles has not received empirical support yet. It is possible that even if learners use non-native-like processing mechanisms, these are employed only under certain circumstances, which would in turn raise the question what factors play a role in inducing these idiosyncratic parsing strategies. These problematic issues highlight the importance of submitting the claims of the SSH to further empirical tests.

3. Event-related potentials and language comprehension

The methodological approach taken in this study was the recording of *event-related brain potentials* (ERPs for short). ERPs refer to changes in scalp-recorded voltages associated in time with some sensory, motor or cognitive event (Kutas & Federmeier, 2000). ERPs are extracted from the much larger EEG signal through signal filtering and through the averaging of multiple segments of EEG containing responses to similar events. The underlying assumption is that the ERP waveform will be similar on each presentation of the stimuli while noise and brain activity that is not synchronized to the event of interest will be random from trial to trial. Through the averaging of a large number of trials, the ERP waveform will

become more and more prominent. ERPs associated with different experimental manipulations are assumed to reflect functionally distinct neurocognitive processes if they show qualitative differences in their timing, polarities (positive or negative going waves relative to waveforms elicited by the other experimental conditions) and spatial distribution across the scalp (Frisch et al., 2004).

3.1 ERPs in native language processing

Research on language comprehension has identified a number of characteristic ERP patterns (components) that are taken to index different language-related processes. The ERP component associated with semantic information processing is the so-called N400: a negative deflection between 250 and 500 ms with a characteristically centro-parietal distribution and a peak around 400 ms after stimulus-onset. The N400 was first observed by Kutas and Hillyard (1980) for semantically inappropriate but syntactically correct sentence-ending words such as *socks* in (4).

(4) He spread to warm bread with socks. (Kutas & Hillyard, 1980, p. 203)

The amplitude of the N400 is sensitive to a variety of manipulations such as the semantic fit between the word and the preceding sentential context, word frequency, vocabulary class (*open-class* versus *closed-class* words) and even context effects above the sentence-level (Kutas & van Petten, 2006). It has been suggested that the N400 indexes the ease or difficulty of semantically integrating an incoming item into prior context (e.g. Friederici, 2002). Alternatively, it is assumed to reflect access of information in semantic memory (Kutas & Federmeier, 2000) or activity in a multimodal long-term memory system (Kutas & Federmeier, 2011). Under these assumptions, context plays a role by modulating activation states of information in long-term memory or by influencing the ease or difficulty with which an item can be retrieved from memory, due to factors such as semantic association and predictability (Stroud & Phillips, 2011).

ERP components typically correlated with syntactic processing are left-anterior negativities (ELAN and LAN) and a late posterior positivity (P600). The early anterior negativity (ELAN) occurring around 100-250 ms has been elicited by phrase structure violations. The ELAN is often more pronounced over the left hemisphere and has been observed in English (Lau, Stroud, Plesch, & Phillips, 2006; Neville, Nicol, Barss, Forster, &

Garrett, 1991), German (e.g. Friederici, Pfeifer, & Hahne, 1993; Hahne & Friederici, 1999; Rossi, Gugler, Hahne, & Friederici, 2005), French (Isel, Hahne, Maess, & Friederici, 2007) and Spanish (Hinojosa, Martin-Loeches, Casado, Munoz, & Rubia, 2003). For example, in Neville et al.'s (1991, p. 156) experimental sentences the phrase structure violation was realized by the fact that a possessive noun cannot combine with a preposition in English:

(5) The scientist criticized Max's of proof the theorem.

Friederici (2002) suggested that the ELAN reflects early automatic phrase structure building processes guided by word category information. Nevertheless, the functional interpretation of the component is controversial since it has been reliably observed only under specific conditions. Lau et al. (2006) argued that one important factor is prediction: a strong expectation concerning the word category of an upcoming item is necessary for phrase structure violations to yield an ELAN. In addition, the component is typically obtained with items whose word category is marked by frequent closed-class morphemes making the formal properties of the word highly characteristic of the unexpected category. Dikker, Rabagliati and Pylkkänen (2009) related this observation to the fact that the ELAN co-occur in time with early sensory processing and argued that the component might reflect the response of sensory areas to an incoming item whose formal characteristics are unexpected on the basis of previously generated syntactic predictions. Herrmann, Maess, Hasting and Friederici's (2009) MEG study showed that Dikker et al.'s (2009) sensory hypothesis applies not only to the visual but also to the auditory domain.

Morphosyntactic violations often elicit a left anterior negativity around 300 and 500 ms (LAN). For example, a LAN has been observed for agreement violations in a variety of languages (Molinaro, Barber, & Carreiras, 2011). In addition, grammatically correct sentences that contain long-distance dependencies also tend to yield an anterior negativity (see Section 6).

The P600 component, a positivity with an onset around 600 ms has been observed with a wide range of violation types such as phrase structure violations (e.g. Friederici, Hahne, & Mecklinger, 1996) and morphosyntactic violations (e.g. Hagoort, Brown, & Groothusen, 1993). Friederici and Frisch (2000) obtained a P600 at the verb with subcategorization violations involving the number of arguments (see (6)): an intransitive verb (*departed*) appeared with two noun phrase arguments (*the inspector* and *the banker*) in the experimental sentences.

(6) Anna weiß, dass der Kommissar den Banker abreiste und wegging.
Anna knows that the inspector the banker departed and left. (Friederici & Frisch, 2000, p. 481)

The presence of an overt syntactic violation is, however, not a prerequisite for the P600: syntactically correct sentences with a non-preferred structure (garden-path sentences) also tend to elicit a late positivity at the disambiguation point where structural reanalysis becomes necessary. For example, Osterhout, Holcomb and Swinney (1994) investigated sentences such as (7): the verb *charge* can be followed either by a direct object or a clausal complement, and NSs have a tendency to initially interpret the postverbal NP (*the defendant*) as a direct object. In this way, the processing of the auxiliary *was*, which makes it clear that the NP is in fact an embedded clause subject, leads to reanalysis.

(7) The lawyer charged the defendant was lying. (Osterhout et al., 1994, p. 790)

The results of some studies suggest that there might be differences in the distributional pattern of the P600 elicited by violations and garden path sentences. Syntactic violations seem to yield a more posterior P600 while structures provoking reanalysis processes result in a more frontally distributed late positivity (e.g. Friederici, Hahne & Saddy, 2002). Finally, sentences containing long-distance dependencies have also been reported to yield a P600 (see Section 6). As far as the functional interpretation of the late positivity is concerned, it has been claimed to reflect syntactic reanalysis and repair processes (Friederici et al., 1996), or more generally syntactic integration cost (e.g. Kaan, Harris, Gibson, & Holcomb, 2000).

Recently, the strict distinction between ERP components indexing semantic or syntactic processing has been questioned by a series of studies that obtained a P600 instead of an expected N400 with certain types of semantically anomalous but syntactically correct sentences (e.g. Bornkessel-Schlesewsky et al., 2011; Kim & Osterhout, 2005; Kolk, Chwilla, van Herten, & Oor, 2003; Kuperberg, Caplan, Sitnikova, Eddy, & Holcomb, 2006; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003; van Herten, Kolk, & Chwilla, 2005). This so-called *semantic P600* has typically been observed in sentences such as (8) in which the anomalous word (*javelin*) would have been a semantically plausible argument of the verb *thrown* with reversed thematic role assignment (so-called *animacy reversal*). In addition, the verb and its argument were often characterized by a close semantic relationship.

(8) The javelin has thrown the athletes. (Hoeks, Stowe, & Doedens, 2004, p. 62)

The functional interpretation of the semantic P600 is still a controversial issue. Nevertheless, the various accounts that have been proposed to explain the phenomenon seem to share the assumption that the semantic P600 generally arises when the relationships that can be established between the verb and its arguments on the basis of the syntactic structure conflict with the most plausible thematic relations among the sentence constituents (Stroud & Phillips, 2011).

3.2 ERPs in L2 processing

Semantic violations elicit an N400 in L2 learners similarly to NSs, but component characteristics might differ across the two groups. Factors that play an important role in this variation seem to be the age of acquisition and proficiency level; however, results obtained by different studies are not consistent in this respect. For example, a delay in onset time was found not only in late learners of the L2 (Weber-Fox & Neville, 1996; age of first exposure after 11) but also in bilinguals who were exposed to their non-dominant language early on in life (Moreno & Kutas, 2005). The amplitude and latency of the N400 might also show variation, and Hahne (2001) observed differences between NSs and learners even in terms of the processing of correct sentences: the amplitude of the N400 elicited by the control items was generally smaller in the latter group. Furthermore, the N400 component in Hahne and Friederici's (2001) study had a longer latency in learners than in NSs suggesting that it might take longer for L2 learners to semantically integrate an incoming word with the preceding context during language comprehension.

Syntax-related components are obtained much less reliably than the N400 in L2 learners. With regard to the ELAN, phrase structure violations failed to yield this component in Weber-Fox and Neville's (1996) experiment, even in learners who started acquiring the L2 as early as age 1-3. Hahne and Friederici (2001) reported similar results with their late and moderately proficient learners. By contrast, advanced late learners in Isel's (2007) and Rossi, Gugler, Friederici and Hahne's (2006) studies displayed an ELAN for phrase structure violations. One possible explanation for the inconsistent results is that the ungrammatical structures in Isel (2007) and Rossi et al. (2006) constituted violations in both the L1s and the

L2s of the participants due to close similarities between the languages, while Weber-Fox and Neville (1996) and Hahne and Friederici (2001) used language pairs that were structurally different from each other (Chinese-English and Japanese-English). In order to investigate the role of structural similarity and language proficiency in the elicitation of syntax-related components, Hahne (2001) replicated Hahne and Friederici's (2001) experiment with highly proficient learners whose first language, Russian, was similar to the L2 in terms of the structures that constituted the violation. Still, no ELAN was obtained and it was suggested that the highly automatic processes that this component potentially reflects might not be attainable for late language learners due to neurological maturational constraints. Nevertheless, results of the above-mentioned studies are not entirely consistent with this assumption as age of acquisition is apparently not always a critical factor, although it might play an important role in interaction with other variables.

The P600 is more frequently reported in L2 studies than the ELAN, while it does not occur as reliably as the N400. For instance, phrase structure violations in Weber-Fox and Neville (1996) led to a P600 only in those learners who started language acquisition before the age of 16. It is important to mention that proficiency and age of acquisition was correlated in this study. Higher proficiency and structural similarity between the languages apparently resulted in an elicitation of the late positivity in Hahne (2001) as opposed to Hahne and Friederici (2001); in Rossi et al. (2006) both moderately and highly proficient learners displayed a P600 for violations that were ungrammatical not only in the L2 but also in the L1. The late positivities observed in these studies were often delayed. Results are again not consistent concerning the factors that contribute to this variation: the P600 elicited in the highly proficient learners of Hahne (2001) had its peak latency as late as 950 ms, while a delay was observed only in the moderately proficient learners and not in the advanced ones in Rossi et al. (2006).

The LAN for morphosyntactic violations also seems to be modulated by factors such as proficiency level and structural similarity (van Hell & Tokowicz, 2010). In detail, proficient learners acquiring an L2 that is structurally similar to their L1 tend to show the same biphasic pattern to morphosyntactic violations as NSs do (LAN followed by a P600). Less proficient learners might only display a P600 (e.g. Rossi et al., 2006). In case of structurally dissimilar languages, high proficiency level seems to be a prerequisite for the LAN to occur (e.g. Ojima, Nakata, & Kakigi, 2005).

In general, the available evidence suggests that the most important factors in the elicitation of the syntax-related ERP components for an L2 are structural similarity and

proficiency level. Note that the apparently significant role of L1-L2 pairings in obtaining native-like ERP responses is not consistent with the predictions of the SSH according to which the parsing strategies that learners use for L2 comprehension are not determined by the L1 background. Nevertheless, the fact that the N400 is much more reliably observed than the syntax-related effects is in accordance with the assumption that learners rely more on semantic information than on structural cues.

A delay in the onset and the peak latency of the N400 and the P600 has been reported with both structurally similar and dissimilar languages and with various proficiency levels. It has been suggested that this delay reflects slower processing of the L2 than the L1 (e.g.: Felser & Clahsen, 2009; Weber-Fox & Neville, 2009). Compared to NSs, learners indeed tend to be slower on lexical decision and grammatical judgement tasks, which might indicate that it takes longer to access the semantics of the L2 words and/or parsing is delayed. In addition, learners' working memory span is often lower in their L2 than in their L1, and it is the L2 working memory capacity that tends to be predictive of L2 comprehension abilities (McDonald, 2006). For example, Service, Simola, Metsanheimo and Maury (2002) suggested that comprehending a non-proficient language might impose an extra burden on working memory as a result of simultaneously using parsing cues (e.g. animacy, word order) from both the L1 and the L2. Consequently, processing speed and working memory load might also play a role in explaining observed differences between native and L2 language comprehension, especially in situations that impose great demands even on NSs' processing resources.

4. Filler-gap dependencies

The present study focuses on the processing of sentences with filler-gap dependencies, in which a phrase appears in a position that is distant from the verb (or other head) that it is an argument of. The canonical position of the phrase in the sentence is called the *gap* and the phrase itself is referred to as the *filler*. The terms *filler* and *gap* are commonly used in psycholinguistics to denote the components of the dependencies without implicating commitment to a specific theoretical account (Phillips & Wagers, 2007), although not all grammatical approaches would model a gap in these constructions. Sentence types containing filler-gap dependencies in English include, among others, *wh*-interrogatives (9-10), relative clauses (11) and topicalization (12) (examples are from Phillips & Wagers, 2007, p. 743):

(9) Which students did the teacher say _____ were falsely accused _____ of the crime.

(10) The teacher said which students _____ were falsely accused _____ of the crime.

(11) The aristocrat hired a young maid who he realized _____ would become his closest confidante.

(12) These chapters, most critics agree you can safely skip _____.

When a filler is encountered during language comprehension, it is not possible to immediately integrate it with the emerging sentence representation; the filler needs to be kept in memory until an appropriate verb that it can be associated with is reached. Different accounts have been proposed to explain how this association is established in real-time. One approach assumes that the mental representations of these sentences involve a structurally present gap, an empty element at the canonical position of the filler; and integration with the verb is mediated by associating the filler with the gap when it is reached during parsing (e.g. Nicol & Swinney, 1989). The semantic compatibility between the verb and its argument is evaluated in addition to assessing whether the structural prerequisites are present for establishing the relationship (e.g. Felser et al., 2012). Alternatively, the *direct association hypothesis* claims that there is no need to postulate structurally present gaps since the filler is directly associated with the argument structure of the verb (e.g. Pickering & Barry, 1991). The term *gap* is used in this paper without commitment to any of these theories, i.e. it refers to the location where the filler phrase would occur if there were no separation between the verb and the position where the filler is actually pronounced, regardless of whether there is an empty category or not at the canonical location of the filler.

Several studies have demonstrated a preference in NSs for completing dependencies as soon as possible: filler-gap relations are constructed immediately when an element that appears to be able to provide the filler with its thematic interpretation is encountered. Language users thus do not seem to await unambiguous information that would confirm the dependency relationship (*active filler hypothesis*: Clifton & Frazier, 1989). One source of evidence comes from the so-called filled gap effect. For example, Stowe (1986) found that in embedded *wh*-questions like (13), reading times were longer for overt direct objects (*us*) that occupied positions where a gap could be expected as compared to the equivalent words in sentences without filler-gap dependences (14). It seems that by the time the readers reached the word *us*, they had already linked the filler to the preceding verb, without waiting for the following items to ensure that the verb does not have an overt direct object.

(13) My brother wanted to know who Ruth will bring us home to at Christmas.

(14) My brother wanted to know if Ruth will bring us home to Mom at Christmas.

The filled gap effect provides information about the time course of filler-verb association, but it does not decide between competing accounts of dependency formation: when an appropriate verb is encountered, it might be either directly associated with the filler or a structural gap might be posited in a following position (Phillips & Wagers, 2007). Proponents of the gap-based theory often refer to the results of cross-modal priming studies that demonstrated filler reactivation at the gap position. These observations were interpreted to reflect the fact that the filler is linked to its canonical structural location. For example, Nicol and Swinney (1989, p. 8) found antecedent-priming effects after the embedded verb (*accused*) in sentences such as (15).

(15) The policeman saw the boy that the crowd at the party accused of the crime.

Nevertheless, since the assumed gap immediately followed the verb in these sentences, reactivation effects seen at this position are in fact compatible with the direct association hypothesis as well. From this perspective, the findings of these studies reflect reactivation of the filler at the verb itself when these elements get directly integrated with each other (Pickering, 2000). Considering these limitations, decisive evidence might be obtained with verb-final languages where all arguments can precede the verb. For example, Aoshima, Phillips and Weinberg (2004) demonstrated filled gap effects at pre-verbal positions in Japanese. In German, Fiebach, Schlesewsky and Friederici (2002) investigated the processing of object *wh*-questions: ERP responses reflecting increased integration costs (P600, see Section 5) were obtained at the case-marked subject noun and not at the clause-final verb. These findings are consistent with the assumption that the filler is linked to a structurally defined gap, which might even precede the verb. Direct semantic association with the verb cannot reasonably take place before the actual element is encountered in the sentence. Still, it is possible to propose a modified version of the direct association hypothesis that could explain these results: if both of them are positioned before the verb, direct association might take place between two arguments, and thematic relationships tentatively established are assessed when the verb is reached (Fiebach et al., 2002). Thus, the psychological reality of gaps continues to be a controversial issue.

4.1 Filler-gap dependency processing in an L2

The SSH makes some specific assumptions concerning learners' on-line construction of filler-gap dependencies. Clahsen and Felser (2006) formulated these predictions in the following way:

> Given the . . . hypothesis that L2 learners underuse syntactic information when processing their L2, it is conceivable that they might try to semantically integrate a displaced constituent directly with its subcategorizer when the latter is encountered, rather than projecting full-fledged grammatical representations that include syntactic gaps. (p. 21)

These claims are more or less consistent with the available empirical evidence on the real-time processing of dependencies in an L2, which will be reviewed below.

Williams, Möbius and Kim (2001) found that learners of English with a range of typologically different L1 backgrounds behaved similarly to NSs in that they employed the active filler strategy to complete dependencies at the earliest possible point. The researchers focused on *wh*-questions with adjunct-extraction such as (16), in which the clause-initial *wh*-phrase was either a plausible (*which girl*) or an implausible (*which river*) direct object of the verb (*push)*.

(16) Which girl/which river did the man push the bike into late last night? (p. 516)

The experiment involved an on-line plausibility judgement task during which reading times were measured. When the *wh*-phrase was a plausible direct object of the verb, both NSs and L2 learners indicated more frequently that the sentence stopped making sense at or after the postverbal noun (*bike*) compared to sentences in which the *wh*-phrase was an implausible direct object. These results were taken to indicate that both groups first interpreted the *wh*-phrase as the object of the verb *push* and were sensitive to the plausibility of this interpretation suggesting that they almost immediately assessed the semantic compatibility of the elements. The initial misinterpretation of the *wh*-phrase as the object led to subsequent reanalysis. Here, some differences between the two groups of participants emerged: plausibility of the *wh*-phrase affected the reading times of NSs already at the determiner *the* while this influence only became apparent on the following noun (*bike*) in case of the learners. The researchers argued the learners were influenced by plausibility to a greater extent than NSs and they took longer to initiate reanalysis.

Juffs (2005) reported further differences between NSs and learners. They investigated subject and object extraction from finite and nonfinite clauses in a self-paced reading study, and the analysis of reading times showed that the learners experienced significant difficulty with subject-extracted finite clauses (17) compared to the corresponding nonfinite (18) or finite object-extracted clauses (19).

(17) Who does the nurse know ____ saw the patient at the hospital?
(18) Who does the boss expect ____ to meet the customers next Monday?
(19) Who does the nurse know the doctor saw ____ in his office? (p. 129)

Reading times in (17) were longest in the region of the embedded verb, and since this sentence type differed from the other conditions in that the following word was also a verb, the observed processing problems were attributed to the presence of two finite verbs next to each other. It is not clear whether these results suggest the existence of qualitative differences in native and L2 processing mechanisms: for example, two appropriate verbs immediately following each other might interfere with the application of a largely semantic strategy for associating the filler with its thematic role assigner. Nevertheless, it is also possible that both groups were garden-pathed by the co-occurring verbs, but learners took longer to recover from the misanalysis.

Felser and Roberts (2007) conducted a cross-modal priming experiment with English indirect object dependencies such as (20), and obtained results that do not support the assumption that L2 learners construct structurally defined gaps at the canonical location of the filler as certain accounts of dependency formation predict for NSs. In this study, participants listened to sentences such as (20), and pictures that were either related or unrelated to the filler word (*squirrel*) were presented at the assumed gap position (after *rules*) and at a control position preceding it (after *game's*). The gap position was not directly next to the verb in these sentences.

(20) Fred chased the squirrel to which the nice monkey explained the game's difficult rules __ in the class last Wednesday. (p. 20)

No antecedent reactivation effects were obtained at the direct object gap position suggesting that the integration between the filler and the verb was not mediated by structural gaps. At the same time, high working memory span NSs in an earlier study displayed these priming effects in the same sentences (Roberts, Marinis, Felser, & Clahsen, 2007), which might

indicate the use of different parsing strategies. Furthermore, learners differed from low working memory span native speakers in that they responded to identical picture targets more quickly than to unrelated ones at both the control and the gap position. L2 learners apparently processed the experimental sentences in the same way regardless of their working memory capacity, and Felser and Roberts (2007) concluded that they were able to keep the filler active in memory. Nevertheless, the results do not necessarily bear on the question if learners can maintain the filler in working memory between the onset and the end of the dependency. Both gaps were located after the subcategorizing verb, around the direct object; therefore, priming effects might simply have reflected compositional semantic integration between the filler, the verb and the direct object.

Felser et al. (2012) used eye-movement recordings to compare the way so-called island constraints affect native and L2 dependency construction. Island constraints refer to the phenomenon that it is not possible to form filler-gap dependences over certain contexts, for example over an intervening *wh*-clause. It seems that islands influence real-time language processing: for instance, they apparently block the operation of the active filler strategy as shown by the absence of filled gap effects within island environments (Phillips & Wagers, 2007). Island constraints have been suggested to derive from general limitations on cognitive processing: referential processing load and additional working memory costs in these structures might create an overload effect (Felser et al., 2012). The experimental sentences of Felser et al. (2012) included structures with (22 and 24) and without island constraints (21 and 23), and the researchers either manipulated the plausibility of the filler as the direct object of the embedded verb (21-22) or they created a filled gap after the verb (23-24). For example, *magazine* in (21-22) was a plausible and *shampoo* an implausible direct object of the verb *read* although the verb did not have an actual direct object NP in these sentences. In (23-24) the direct object position of *read* was filled by a noun phrase (*articles*).

(21) Everyone liked the (magazine) shampoo that the hairdresser read extensively and with such enormous enthusiasm about before going to the salon.

(22) Everyone liked the (magazine) shampoo that the hairdresser who read extensively and with such enormous enthusiasm bought before going to the salon. (p. 77)

(23) Everyone liked the magazine that the hairdresser read articles with such strong conclusions about before going to the beauty salon.

(24) Everyone liked the magazine that the hairdresser who read articles with such strong conclusions bought before going to the beauty salon. (p. 87)

The plausibility manipulation had an effect on NSs' and L2 learners' reading times at the verb *read* in nonisland environments, but they showed no attempt at dependency formation between the filler and the verb within the island environments: both NSs and L2 learners were apparently influenced by island constraints during real-time language comprehension. In other respects, however, the two groups behaved differently: learners displayed immediate sensitivity to plausibility information on dependency formation (i.e. in their first-pass reading times of the verb *read* and the following word) but their reaction to the information provided by the filled gap became apparent only later, possibly only at a reanalysis stage (i.e. in their rereading times of the next three words after *articles*). The opposite pattern was observed with NSs. Felser et al. (2012) argued that their findings are compatible with the SSH since, unlike NSs, learners seemed to assess the semantic compatibility of the filler and the verb instantly while syntactic information concerning the possibility of dependency formation was evaluated at a later stage. Furthermore, they tentatively suggested: "(certain parts of the) L2 grammar knowledge may be represented in such a way so as to make it inaccessible to first-pass parsing routines, and available only during later stages of processing" (p. 95). From this perspective, L2 learners opt for the strategy of shallow parsing because it enables faster processing than full parsing does. The reason why non-native grammar knowledge would be represented in the proposed way was not discussed.

To conclude this section, learners' processing of filler-gap constructions appears to be similar to native language comprehension in that they show a preference for completing dependencies as early as possible while they are also sensitive to island constraints. These findings are not surprising given that both of these phenomena have been associated with limitations on general cognitive processing. A related issue concerns learners' ability to maintain the filler in working memory throughout a dependency, but the available evidence does not allow us to draw any firm conclusions. With regard to the question whether learners directly integrate the filler with the verb as predicted by the SSH, the results are again inconclusive. Although learners appear to process these structures differently from native speakers, it is not clear where exactly the difference lies: Williams et al. (2001) and Juffs (2005) do not provide evidence directly bearing on the question of how the filler is linked to the verb. Felser and Roberts's (2007) results only suggest what learners do *not* do in order to be able to complete the dependency. Felser et al. (2012) obtained findings on the relative timing of the influence of presumably different information sources: it seems that, at least initially, a semantic integration takes place between the filler and the verb. However, syntactic information also appeared to have an effect, which does not unambiguously support

the SSH. Nevertheless, it is not clear whether learners' reaction to the filled gap manipulation indeed reflected sensitivity to structural cues (e.g. recognizing that the possible structural position of the filler is occupied by another item) or they simply had difficulties with semantically integrating an additional and unexpected noun. To clarify these issues, it is necessary to gain more fine-grained information on the timing and nature of different types of processes that play a role in language comprehension. The event-related potential technique enables us to obtain such information.

5. ERPs and filler-gap dependencies in native language processing

One ERP component that typically accompanies native speakers' processing of filler-gap dependencies is an anterior negativity than tends to be larger over the left hemisphere. The negative effect starts at the filler phrase or at the element directly following it around 200-400 ms after word onset, and it has also been observed at the point where the filler-gap dependency can be completed (e.g. Kluender & Kutas, 1993; Felser, Clahsen, & Münte, 2003). For example, Kluender and Kutas (1993) reported an anterior negativity at the word *he* in sentences like (25) and also at the word *into* in sentences like (26).

(25) Have you forgotten who he dragged ___ to the movie that weekend? (p. 200)

(26) Can't you tell what she intends to drum ___ into you by the end of the quarter? (p. 203)

Studies that computed ERPs for regions spanning multiple words demonstrated that the negativity continues throughout the dependency (e.g. Fiebach et al., 2002; Phillips, Kazanina, & Abada, 2005). Consequently, the effect is often referred to as the sustained anterior negativity (abbreviated as SAN). The SAN is generally interpreted to index the cost of keeping the filler in working memory (WM for short) as well as the subsequent retrieval and/or reactivation of the filler when it can be associated with its subcategorizer. Consistently with this assumption, the SAN appears to be sensitive to length manipulation: Fiebach et al. (2002) reported that as the distance from the filler grew, the amplitude of the negativity increased. It also typically, but not exclusively, appears in long-distance dependencies with several intervening words between the filler and the subcategorizer. In case it is observed in shorter dependencies, the latency of the effect might also be significantly shorter presumably indexing lower WM demands. In addition, the component seems to be modulated by WM

span differences: its distribution was found to be broader in individuals who had a low WM span (Fiebach et al., 2002). In general, the assumption that anterior negativities can index WM costs was supported by Vos, Gunter, Kolk and Mulder's (2001) study that investigated the nature of interaction between the processing of morphosyntactic violations and WM capacity: additional WM load elicited sustained frontal negativities throughout the length of the experimental sentences, and the latency of this negativity was influenced by individual WM span.

WM load induced by the filler is probably reflected in the SAN; however, it is not clear exactly what is maintained in WM. Fiebach et al. (2002) enumerated the possibilities: the maintained information might be the filler as such, only certain features of the filler, or predictions concerning the constituents that are necessary to form a grammatical clause (Gibson, 1998). Fiebach, Schlesewsky and Friederici (2001) tentatively suggested that only syntactic information is maintained and Wagers and Phillips (2012) obtained results that are consistent with this assumption. Based on the findings of a reading time study, they argued that only the word category of the filler is kept in the focus of attention to enable the generation of structural predictions based on the previous input. At the same time, this limited amount of information kept active in memory makes it possible to allocate resources for the processing of intervening material between the filler and its subcategorizer, and when the dependency can be completed, unmaintained features of the filler need to be retrieved.

The processing of the word constituting the end point of the dependency tends to elicit a late posterior positivity compared to the equivalent word in sentences that do not contain a dependency (Felser et al., 2003; Fiebach et al., 2002; Kaan et al., 2000; Phillips et al., 2005). This positivity typically appears at the subcategorizing verb and it is similar to the P600 for syntactic anomalies concerning its distribution and timing. Unlike in the case of the SAN, the amplitude of the P600 is not influenced by WM capacity; therefore, it does not appear to index WM load (Fiebach et al., 2002). Instead, it is generally interpreted to reflect syntactic integration cost (e.g. Kaan et al., 2000), which tends to be higher at the subcategorizing verb in a dependency than in the control sentences. For example, Kaan et al. (2000) obtained a P600 at the verb *imitated* in sentences such as (27) relative to the control condition (28). Kaan et al. argued that at the point of processing the verb *imitated*, integration takes place between the verb and the subject *performer in the concert* in both sentences, whereas in (27), the *wh*-phrase *who* is also integrated as the object leading to higher integration costs in this condition.

(27) Emily wondered who the performer in the concert had imitated for the audience's amusement.

(28) Emily wondered whether the performer in the concert had imitated a pop star for the audience's amusement. (p. 164)

Fiebach et al. (2002) also argued that the amplitude of the P600 depends on the number and difficulty of integrations computed at the verb, and Phillips et al. (2005) emphasized the role of assessing the thematic relationships between the verb and the filler, subsequent compositional interpretation of these elements and, in general, the number of thematic role assignments taking place at the verb.

5.1 Hestvik et al.'s study

As it was discussed in Section 4 the presence or absence of structural gaps in NSs' mental representations of filler-gap dependencies is still a controversial issue. Hestvik et al. (2007) conducted an ERP experiment with native English speakers to evaluate competing theories of filler-verb association. In their experimental sentences, the assumed gap position, i.e. the canonical position of the sentence-initial object noun, was ungrammatically filled by an overt noun phrase (*the camel in* (29)):

(29) The zebra that the hippo kissed <u>the camel</u> on the nose ran far away.

(30) The zebra said that the hippo kissed the camel on the nose and then ran far away.

(31) The zebra that the hippo kissed on the nose ran far away. (p. 303)

The direct association hypothesis predicts that the extra noun after the verb would create processing problems that are semantic in nature such as providing this noun with a thematic role. Consequently, an N400 is expected at this position. Hestvik et al. (2007), however, found that the extra NP yielded an ELAN component relative to the control conditions (30-31) suggesting that the filled gap constituted a phrase structure violation for the participants.

Hestvik et al. (2007) argued that their findings support theories that would predict the construction of a structurally present gap after the verb *kissed.* They adopted Friederici's (2002) model of auditory sentence comprehension to provide a basis for the interpretation of the observed components. In this model, the ELAN is assumed to reflect an early phase of sentence processing when an initial structure is built up on the basis of word category information without taking argument structure and semantic well-formedness into account. In detail, the possible minimal structures of a language are supposed to be learnt and stored as

templates, which are quickly activated when the category of an incoming word is accessed. The failure to match the input with a template constitutes a phrase structure violation (Friederici, 2011). According to Hestvik et al.'s argument, encountering the extra noun phrase (*the camel*) violates word category expectations only if the parser postulates a structural gap with the same properties as the sentence-initial noun phrase (*the zebra*) after the verb *kissed*. Without an actually present gap, the NP V NP sequence (*the hippo kissed the camel)* is a well-formed relative clause structure in English (e.g. *The book that John gave Mary…*) whereas an additional gap position would create an ill-formed relative clause: NP V NP NP (e.g. *The book that John gave Mary the present…*). Note that the observation of the ELAN in these sentences is consistent with the sensory ELAN hypothesis as well: the word *the* is a closed-class morpheme and it is formally highly typical of the category that creates a phrase structure violation in this position.

No N400 was obtained after the ELAN in this experiment, which is consistent with the results of studies that investigated the effects of combined word category and semantic violations and found that these conditions yield only an ELAN. According to Friederici (2002), the identification of a phrase structure error during initial syntactic structure building prevents further semantic and syntactic processing.

6. The present study

Hestvik et al.'s (2007) experiment demonstrated how ERPs can be used to distinguish between the predictions of competing theories on filler-gap dependency processing. Recall that the SSH assumes that L2 learners semantically associate the filler with the thematic structure of the verb, and this process is not mediated by structural gaps. With regard to Hestvik et al.'s experimental sentences, the SSH predicts that the filled gap would elicit an N400 in L2 learners of English and not an ELAN: the N400 would reflect that learners try to integrate the extra argument semantically, and they experience problems with interpreting the semantic role of the additional NP. Obtaining an ELAN effect would challenge the SSH as this result would imply that the establishment of an initial link between the filler and the verb involves the use of cues provided by phrase structure representation.

Hestvik et al.'s (2007) design enables us to make clear predictions concerning the implications of the obtained ERP effects for the SSH. The present study adopted the three experimental conditions of Hestvik et al. (see 29-31) to investigate the assumptions of the

SSH that L2 learners establish a direct semantic association between the filler and the verb without the reliance on syntactic cues. In addition, further conditions were introduced to be able to gain more information on the way learners process filler-gap dependencies in English relative clauses. ERP responses to the word *that* and to the embedded verb were also compared across the different conditions to investigate issues concerning the maintenance of the filler in WM and the integration processes taking place at the verb. The results were evaluated against the findings of previous ERP studies on native dependency processing. Examples for four conditions of the study are shown below; (32), (33) and (34) correspond to the three experimental conditions of Hestvik et al.:

(32) The writer that the assistant greeted ____ at the meeting sat down.

(33) The writer that the assistant greeted <u>the teacher</u> at the meeting sat down.

(34) The writer said that the assistant greeted the teacher at the meeting and then sat down.

(35) The writer that ____ greeted the assistant at the meeting sat down.

Sentence (32), (33) and (35) involve a filler-gap dependency. The term *filler* is used in this study to denote the noun phrase (*the writer*) that appears in a position that is different from its canonical location as the object of the verb in (32) and (33) or as the subject of the verb in (35). In this study, *filler* refers to the full semantic and syntactic content of the noun phrase, and the following relative pronoun *that* is assumed to be co-referential with the noun phrase. The term *gap* refers to the point in the linear sequence of words where the filler would be pronounced if it occurred in a main clause or non-relativized subclause, e.g. the slot following the embedded verb *greeted* in (32) and (33) and the position after the relative pronoun *that* in (35). The noun phrase (*the teacher*) occupying the gap slot in (33) is not regarded as a filler in the above-described sense; it is instead referred to as an extra argument or object.

Following Clahsen and Felser (2006), a distinction is made between syntactic and semantic information types, while it is acknowledged that only certain grammars assume these distinctions. Syntactic information potentially employed during sentence comprehension is assumed to derive from the syntactic structure of an utterance involving hierarchical organization of words into units on the basis of permitted combinations of word categories. Word category distinguishes between groups of words based on their

morphosyntactic properties. In addition, a distinction is made between the syntactic and semantic constraints that a subcategorizer (a verb or other lexical head) imposes on its arguments. Syntactic requirements concern the way the arguments of the subcategorizer are syntactically realized in terms of their number and word category whereas semantic restrictions involve the thematic roles that the arguments play. The experimental conditions of the present study are presented in detail after a discussion of some issues around the selection of the participants.

6.1 Implications for the selection of the participants

The review of previous ERP research on L2 comprehension (Section 3.2) has highlighted a number of factors that can potentially affect the way learners process their L2; the most important ones seem to be language proficiency, L1-L2 similarity and age of acquisition. Consequently, it was necessary to take these factors into consideration during the selection of the participants.

The participants in the study were advanced L2 learners of English with Swedish as a native language. Swedish and English are typologically close to each other, both being Germanic languages. The structures investigated are also formed similarly in these two languages (see (36) for a Swedish object-relative clause and its English translation): for example, a relative pronoun (*som*) introduces the object-relative clause and the object noun (*Den bok*) appears clause-initially. In the canonical position of the object (after *köpte*), there is no element such as a pronoun that would refer back to the noun phrase.

(36) Den bok som Anna köpte ____ igår ligger på bordet.
The book that Anna bought ____ yesterday is lying on the table.

The selection of the L1-L2 pair and the proficiency level of the participants were based on the following considerations: ERP responses obtained in L2 learners in the present study are compared to findings previously reported with NSs. As discussed in Section 3.2, native-like syntax-related ERP components were typically reported with proficient learners and when the L1 and the L2 were similar concerning the violated structures. Therefore, previous findings suggest that these ERP effects can at least potentially be observed with advanced Swedish learners of English. In this way, a possible failure to obtain native-like ERP components would still enable us to draw conclusions concerning the implications of the findings for the SSH. Note that the SSH claims that L2 learners employ shallow parsing strategies regardless

of the similarities and differences between the L1 and the L2. By choosing advanced participants we can also exclude the possibility that non-native-like performance would simply reflect learner's incomplete grammatical knowledge of the examined sentence types.

With regard to the age issue, it is a general observation that there is a relationship between age and the outcome of L2 acquisition; however, the exact role that age plays in the success of language learning is a controversial issue (Munoz & Singleton, 2011). Some researchers argue that there are maturational constraints on L2 acquisition (e.g. DeKeyser, 2000; Johnson & Newport, 1989), while others claim that there are no qualitative differences between language learning in childhood and adulthood (e.g. Bialystok & Hakuta 1999; Flege 1999). Instead, observed differences can be attributed to a variety of other factors that are often correlated with age: for example, motivation, quality of received input and the degree of social and psychological integration with the non-native culture. According to the proponents of the maturational account, there is a critical period for language acquisition: a "window of opportunity" characterized by an enhanced language learning ability and by the possibility of achieving native-like proficiency in an L2 (Birdsong, 2006). The offset of the critical period during or at the end of childhood is marked by some maturational event that leads to qualitative changes in language acquisition (Munoz & Singleton, 2011). The critical period has been proposed to selectively or uniformly affect a variety of linguistic domains and skills such as phonology, morphosyntax and implicit learning mechanisms. There is also a wide range of views concerning the age that marks the end of the period: for phonetics and phonology sometime between the age of 1-12 and for syntax around age 4-15 .The exact nature of the maturational changes that could account for the critical period is also far from uncontroversial: neurobiological explanations refer, for example, to the lateralization of language functions, maturation of brain cells and myelination processes in the language areas (Singleton, 2005). It follows from the uncertainty surrounding the potential causes of the critical period that it is also unclear whether maturational constraints would only affect the way linguistic knowledge is represented in the brain or it would lead to non-native processing mechanisms as well.

The concept of the critical period is, therefore, unclear and problematic, which could also have contributed to the fact that no conclusive evidence has so far been provided for or against the hypothesis. Consequently, implications of a potential critical period on language acquisition have to be taken into consideration during the selection of the participants: they should all belong to the same population in terms of the age-factor in language acquisition. The SSH is a hypothesis about late learners sentence processing mechanisms (e.g. Felser et

al., 2012). From the perspective of the critical period it means that the SSH applies to individuals who did not acquire the L2 within the maturational "window of opportunity". Nevertheless, there is an additional factor that needs to be considered: the context of language learning. Studies investigating issues around the critical period hypothesis have predominantly involved learners who were immersed in the L2, i.e. they learned the language in an environment where it was used as an L1. It has been argued that when language learning is restricted to the typical classroom setting, the existence of maturational constraints might be irrelevant as pre-puberty learners cannot take advantage of the early start in the same way as it is possible in the case of immersion (e.g. Johnson & Newport, 1989). Indeed, several of those studies that Clahsen and Felser (2006) based their hypothesis on or that were later conducted to test the SSH involved participants that started L2 learning in a classroom setting around the age of 6-14 but were not immersed in the L2 before adulthood.

As English language teaching begins sometime during the first years of compulsory education in Sweden, it would have been extremely difficult to find Swedish native speakers who had not had any contact with English before adulthood. Probably these individuals would also have been highly unrepresentative of the population of Swedish L2 learners of English. Consequently, and in accordance with the practice followed by Clahsen, Felser and their colleagues, participants who started instructed English language learning in a formal school setting after the age of 6 were allowed to take place in the study. At the same time, it was a requirement that none of the participants should have been immersed in English for a longer period of time (more than 4 weeks) before the age of 16 (the critical period for syntax acquisition is generally assumed to come to an end sometimes before this age).

6.2 Experimental design

Table 1

Example sentences for each of the six experimental conditions

a)	Filled gap The receptionist that the painter scared the reporter by accident answered the phone.
b)	Grammatical gap The receptionist that the painter scared by accident answered the phone.
c)	Grammatical object The receptionist said that the painter scared the reporter by accident and then answered the phone.
d)	Semantic anomaly The receptionist said that the painter scared the document by accident and then answered the phone.
e)	Correct subject-relative The receptionist that scared the painter by accident answered the phone.
f)	Semantically anomalous subject-relative The receptionist that scared the freezer by accident answered the phone.

The filled gap effect was tested through comparisons between ERP responses obtained at the words directly following the verb in the embedded clause in *filled gap, grammatical gap* and *grammatical object* conditions. In the filled gap condition, the object position of the relative clause is ungrammatically occupied by an overt noun phrase. The grammatical object serves as a control condition: it contains the same string of words as the relative clause part of filled gap, but it is grammatically correct due to the fact that the string is included in a complement clause introduced by *that*. The grammatical gap condition consists of grammatical sentences with an object-relative clause in their subject position. Grammatical gap was added as a second control condition as a result of the following considerations: a number of studies have reported a sustained anterior negativity between the filler and the gap (see Section 6) suggesting that sentences of the filled gap condition might elicit a negativity related to increased working memory load whereas this effect is not expected in the grammatical object condition. If comparisons are made only between these two conditions, it might not be possible to decide whether an enhanced negativity obtained at the extra object position in the filled gap sentences is an ELAN, an N400 or only part of a long negativity. The inclusion of grammatically correct sentences with filler-gap dependencies enables us to decide whether filled gaps actually elicit an enhanced negativity in addition to a potential sustained negativity.

The SSH predicts that the extra NP argument in the filled gap condition yields an N400, which would suggest that L2 learners experience problems with semantic integration when processing these structures. This explanation, however, would only be valid if we can demonstrate that the learners of the present study display an N400 also for grammatical sentences containing a clear semantic anomaly of the kind that has been observed to yield this component in NSs and learners. Recording the participants' reactions to obvious semantic violations would also enable us to provide a more valid interpretation of the potential absence of an N400 in the filled gap condition. As a result of the above considerations, a fourth experimental condition was included: the *semantic anomaly* condition consists of grammatically correct sentences that are identical to those in the grammatical object condition, except for the fact that the object NP of the complement clause was replaced with an inanimate noun creating semantically odd sentences.

6.2.1 Subject-relative clauses

Two subject-relative clause conditions were also included in order to increase the percentage of grammatically correct sentences. Hestvik et al. (2009) suggested that the ELAN might be sensitive to the proportion of sentences with phrase structure violation relative to grammatical sentences. In their study, raising the percentage of ungrammatical structures from 15% to 25% seemed to have created an expectation of the violation and led to an attenuation of the ELAN component.

The sentences in the *correct subject-relative* condition are the subject-relative clause equivalents of the sentences in the grammatical gap condition. The *semantically anomalous subject-relative* sentences have the same structure as the items in correct subject relative clause condition, but they contain a semantic anomaly at the object NP in the same way as the sentences of the semantic anomaly condition. Although the subject-relative clause conditions were originally introduced to increase the number of grammatically correct sentences in the stimulus material, including these two conditions in the ERP data analysis enables us to compare the processing of subject-relative clauses and object-relative clauses. Therefore, ERP responses to critical words will also be analysed for the subject-relative clause items, and if comparisons with the equivalent elements in the other experimental sentences yield results that can contribute to the interpretation of the findings of the study, the outcome of these analysis will also be reported and discussed.

ERP responses to the words that mark the onset and the end of the dependency, i.e. *that* and the embedded verb respectively, will also be analysed to gain more information on the way L2 learners comprehend these structures. The subject-relative clause sentences will also be included in the analysis as they provide useful points of comparisons at the above-mentioned words. As sentence (a) and (b), (c) and (d), (e) and (f) are completely identical until after the word *that* and also until after the embedded verb, each sentence pair will be treated together as a single condition during the analysis of the results, which will double the number of samples in each condition and will, consequently, increase statistical power.

6.2.2 *That* condition

In filled gap (a) and grammatical gap (b), *that* is part of a relative clause in which the object of the embedded verb appears sentence-initially. The two sentence types together make up the *object-relative that* condition. The two subject-relative clauses sentences (e) and (f) create the *subject-relative that* condition. The *that* in grammatical object (c) and semantic anomaly (d) introduces a finite complement clause without any filler-gap dependencies. These sentences thus constitute the *complementizer that* condition.

6.2.3 Verb condition

Focusing at the embedded verb position in each sentence, grammatical gap and filled gap together constitute the *object-relative verb* condition. Grammatical object and semantic anomaly creates the *complement clause verb* condition. Finally, the two subject-relative clause sentence types make up the *subject-relative verb* condition.

6.3 Hypotheses

6.3.1 Processing the filled gap

Based on the assumption that L2 learners employ shallow parsing, it is predicted that the filled gap will yield an N400 relative to the two control conditions. If L2 learners processing sentences such as (a) do not construct representations with a gap after the verb *(scared*) but provide the sentence-initial noun phrase *(the receptionist)* with its thematic role as soon as they encounter and access the semantics of the verb, the filled gap will not produce a violation of word category expectations and no ELAN will be obtained. However, by the

time the word occupying the gap position is reached (*the reporter*), both arguments of the verb have been integrated into its thematic structure; therefore, it is not possible to interpret the thematic role of the extra noun phrase. This disruption in the representation of semantic relationships is expected to yield an N400.

Some previous studies with NSs observed responses to extra noun phrases (*the violinist* in (37)) not permitted by the argument structure of the verb (*dawdle*). The experimental sentences did not involve filler-gap dependencies, and, therefore, the extra noun phrase did not constitute a general phrase structure violation as the filled gap did in Hestvik et el.'s (2007) study. Instead, these structures were assumed to pose, at least in part, similar processing problems that the L2 learners of the present study are predicted to face while listening to the filled gap constructions: the integration of an unexpected argument. Friederici and Frisch (2000) and Frisch et al. (2004) indeed found an N400 for these violation types.

(37) Heute trödelte (V) der Cousin (NOM) den Geiger (ACC) am Aufzug.
Today dawdled (V) the cousin (NOM) the violinist (ACC) at the lift. (Friederici & Frisch, 2000, p. 490)

The results of these studies also suggest that the N400 component might be followed by a small P600 effect indexing structural revisions in order to deal with the ungrammaticality caused by the extra argument. The SSH does not necessarily predict a P600 as learners might not even build a detailed enough structure that would require reanalysis. Nevertheless, if we consider Felser et al.'s (2012) proposal that some part of L2 learners' grammatical knowledge might become available at a later stage of processing, a late component such as the P600 might be obtained reflecting the evaluation of the representations that have been built up from the point of view of grammaticality.

Both the N400 and the P600 effect have been reported to be delayed or reduced in L2 learners, possibly even in highly proficient ones (see Section 3.2). These components are, therefore, predicted to occur later than what would be expected on the basis of studies with native speakers. Finally, the object NP in the semantic anomaly sentences (*the document*), which were added to enable the interpretation of the predicted effect for the filled gap, is expected to yield an N400 relative to the equivalent NP in grammatical gap sentences (*the reporter*) reflecting difficulties with semantic integration or with semantic access. Since the subject-relative clauses involve the same semantically correct versus semantically anomalous word pairs following the embedded verb as the semantic anomaly and grammatical gap

sentences, the object NP in the semantically anomalous subject-relative clauses (*the freezer*) is predicted to elicit a similar N400 relative to the object NP in the correct subject-relative clauses (*the painter*).

6.3.2 Processing filler-gap dependencies - *That*

As discussed in Section 6, a sustained anterior negativity (SAN) that starts around the filler phrase and lasts until the point where the dependency can be completed has been observed in NSs, and this component has been associated with taking in and retrieving the filler from working memory. Since the maintenance of information is generally expected to elicit working memory costs in learners similarly to NSs, the object-relative that is predicted to yield a sustained anterior negativity in comparison with the complementizer *that*. It is not possible to make clear predictions concerning the timing of this component since previous studies have reported a range of different latencies with the earliest onset time around 200 ms (Martin-Loeches, Munoz, Casado, Melcon, & Fernandez-Frias, 2005). The subject-relative that is not expected to elicit a similar SAN since it is immediately followed by the subcategorizing verb, which makes it unnecessary to actively maintain the filler in working memory.

In the experimental sentences of this study, the number of intervening elements between the filler and the subcategorizing verb is considerably smaller than in most studies that have found a SAN. Consequently, it is also possible that the object-relative *that* will not generate a high enough working memory load for the negativity to be observed. At the same time, the SAN seems to be modulated by differences in working memory span (Fiebach et al., 2002), and given the assumption that the processing of an L2 requires greater working memory resources than the processing of an L1, a SAN might appear during L2 comprehension even in cases when it is absent in NSs. Moreover, the present study will use auditory stimuli while most other ERP experiments investigating this topic have employed visual presentation. Due to additional processing costs associated with understanding auditory stimuli such as phonological decoding, L2 learners might find the comprehension of spoken language more demanding than that of written language, and the increased working memory load might be reflected in a SAN.

Alternatively, comprehending spoken sentences in an L2 might in itself be so demanding that it is not possible to purposely allocate resources to maintain the filler in working memory. It seems that even NSs might keep only a subset of all the features of the

filler in the focus of attention throughout a long dependency (Wagers & Phillips, 2012). If learners did not maintain any features at all, the filler would need to be reactivated upon encountering its potential subcategorizer leading to a negativity appearing only at the embedded verb of the object-relative clause.

6.3.3 Processing filler-gap dependencies - Verb

The component typically observed at the subcategorizing verb in filler-gap dependences is P600-like positivity (see Section 5). As the amplitude of this component has been correlated with syntactic integration cost, which is assumed to depend on factors such as the number of thematic role assignments and the difficulty of the integration processes (Fiebach et al. 2002), an enhanced P600 is predicted for the object-relative verb relative to the complement clause verb and the subject-relative verb. In detail, encountering the embedded verb in all conditions enables its integration with the preceding subject, and the evaluation of the properties of the verb presumably generates an expectation for an upcoming object. In case of the object-relative verb sentences, the object has appeared already at the beginning of the sentence, and, therefore, this additional argument can also be integrated leading to greater processing costs in this condition. At this point, the on-going maintenance or reactivation of the filler in working memory might also lead to an additional anterior negativity (the SAN) in the object-relative verb sentences. Finally, integration processes in the subject-relative verb condition might also be somewhat more difficult than in the complement clause verb condition: in case of the subject-relative verb, the subject noun is not adjacent to its subcategorizer due to the intervening *that*. However, as this distance is very short and the word *that* is assumed to have the same referent as the noun, no significant difference in the amplitude of the P600 is expected across these conditions.

7. Method

7.1 Participants

Fourteen right-handed native speakers of Swedish participated in the experiment (5 women, mean age: 23, range 20-27 years). Three additional participants were excluded from the final data analysis due to excessive EEG artefacts such as blinks and head movements. The subjects were recruited from among the students of Lund University, and they received a cinema ticket for their contribution. On average, the participants started instructed English

language learning in a classroom setting at the age of 9.1 (minimum: 7, maximum: 11), and none of them spent more than 4 weeks in an English-speaking country before the age of 16. The participants rated themselves as advanced but not native-like speakers of English (average ratings on a scale between 1 (low) and 6 (high): speaking – 4.92; listening - 5.21; writing – 4.93; reading - 5.36; grammar – 4.93). English was always the second most proficient language of the participants with Swedish being the first language. Informal observation before the experiment confirmed that all participants spoke English fluently and understood the language without any difficulties.

7.2 Materials

In this experiment, 6 different sentence types were used (see Table 1). The items in conditions (a), (b) and (c) were structurally identical to the experimental sentences of Hestvik et al. (2007) while they differed in terms of the lexical items. Sentences such as (30) might have sounded semantically odd for the adult participants of the present study since Hestvik et al. adopted their stimulus material from an experiment with pre-school children (Love & Swinney, 1997), although their participants were also adults. The animal nouns were replaced with profession nouns, and the rest of the sentence was modified to suit the context set by the nouns. In this way, all arguments of the embedded clause verb were still animate and relatively frequent in spoken language so that they were not expected to pose problems to the learner group.

Forty sets of sentences were created resulting in a total of 240 trials (for a list of the experimental sentences see Appendix). Sentences within each condition were of equal length in terms of the number of syllables until after the last critical word (the item following the embedded clause verb). Each embedded clause verb was used twice but with different arguments, and all of them were transitive verbs that take only one object. Each profession name was also used twice with every profession occurring at least once as a grammatical argument. The verb in the grammatical gap condition was always followed by a prepositional phrase. In this way, comparisons in this position involved only closed-class words (a determiner or a preposition) to avoid effects related to differences in processing closed-class items versus content words.

In the semantic anomaly condition and the semantically anomalous subject-relative condition, the noun phrase following the embedded verb was semantically incongruous with the rest of the sentence. As the sentence context created an expectation for an animate noun

in this position, the semantic anomaly was introduced by replacing the profession name with an inanimate noun that had the same number of syllables. The inanimate nouns were always concrete, except for two cases in which a more abstract noun was used, as frequent concrete nouns in these sentences would not have appeared semantically odd. The same set of inanimate nouns was used in the two semantically anomalous conditions.

In order to make sure that the participants pay attention to the sentences, written comprehension questions were constructed for all items. During the stimulus presentation, only one-fourth of the sentences were followed by a comprehension question. In this way, the participants always had to be prepared for answering a potentially appearing question while the experimental sessions did not become unreasonably long. The questions that were actually used during the experiment were randomly selected from among all the items created for a given condition. The question types were equally distributed across conditions, and the proportion of correct yes and no answers was 50-50%.

The questions focused on the content of the sentences, and in 18 cases on the actual words appearing in the sentences (e.g. *Did you hear the word blueberry?*). Half of the questions were *yes/no*-questions (e.g. *Did the architect spill his drink?*) and the other half were *wh*-questions, which required the selection of one of two alternative answers (e.g. *Who went into the store: the dentist or the photographer?*). In case of the filled gap items, the questions focused on the grammatical part of the sentence.

The sentences were spoken by a male American English speaker in an anechoic room and were recorded digitally with 16 bit at a sampling rate of 44 kHz. After the recordings the sentences were normalized in amplitude. To compensate for the potentially unusual prosody of the ungrammatical extra argument in the filled gap condition, every second item in this condition was constructed by a splicing procedure using the speech editing software Praat. The part of the sentence that contained the extra noun phrase was edited out and was inserted into the equivalent grammatical gap sentence. The beginning and the end point of the removed material were never the same as the positions to which the ERPs were time-locked. In order to ensure that the edited sentences would sound natural, appropriate splicing points were identified by auditory inspection of the relevant portions of the recordings and by visual inspection of the oscillograms and the spectrograms. The same splicing treatment was used to construct every second grammatically correct sentence as well to avoid that the participants' potential reaction to the audio-edited material would be interpreted as a condition effect.

During the stimulus presentation, the sentences were distributed across 5 blocks. Each of the 5 blocks constituted 6 randomly selected sentences from each condition without

questions and 2 randomly selected sentences from each condition with questions. Each sentence was presented only once during the experimental session.

7.3 Procedure

The participants were seated in front of a computer screen and the sentences were presented auditorily in random order via loudspeakers. Before the experiment, the participants were instructed to minimize eye and body movements during the acoustic presentation of the sentences. The stimulus presentation was controlled through a PC using E-prime software.

The experimental session lasted for about 50 minutes and was preceded by a practice block. The practice block consisted of five trials that were similar but not identical to the experimental sentences, and immediate written feedback was provided concerning the accuracy of the participants' responses. Before the presentation of the first experimental sentence, the participants had the opportunity to ask questions for clarification. The experimental session was subdivided into five blocks of 48 trials with short rest breaks in between. First, a fixation point appeared on the screen for 750 ms before sentence onset and remained there during the presentation of the sentence to help reduce eye movements. After the auditory presentation lasting for approximately 5000 ms, the fixation point was still visible for 500 ms. A comprehension question was then displayed on the monitor in case a question was selected for the given trial. The subjects indicated their answer by pressing one of two buttons on a Serial Response Box within 20000 ms. The trial continued as soon as the participants pressed a button, and no feedback was provided concerning their answers. Consecutive trials were separated by an interval of 1000 ms with black screen.

7.3 Data acquisition

The EEG was recorded using a 128 channel HydroCel Geodesic Sensor Net with vertex on-line reference. The channels were re-referenced offline to the average of the 128 recording sites. Data were sampled at a rate of 250 Hz, and an on-line bandpass filter with cutoff frequencies of 0.01 and 70 Hz was used. The impedance of all electrodes was kept below 50 kΩ, as recommended by the manufacturer.

7.4 ERP data analysis

ERP data processing was conducted using Netstation software. Offline, the EEG was

highpass filtered (0.1 Hz) and lowpass filtered (30Hz) in order to reduce noise while keeping those frequencies that make up the most relevant part of the ERP waveform. The continuous EEG was segmented into 1000 ms epochs that were time-locked to the onset of the word *that*, the embedded clause verb, and the first word following the embedded clause verb in each sentence. A time period of 100 ms preceding the epochs was used for baseline correction. ERP measurements are based on the assumption that the variation of the EEG activity before stimulus onset is randomly distributed over conditions; therefore, data from the baseline interval is used to establish the zero-voltage value.

A segment was removed from analysis if it contained more than 10 bad channels, an eye-blink or an eye-movement. A channel was marked as bad if the average amplitude exceeded 200 μV, and it was considered bad in all trials if it was bad in at least 20% of trials. In case of subjects with frequent eye movements (8 subjects), these artefacts were corrected for by submitting the data to a procedure called Ocular Artefact Removal available in Netstation. In this way, a larger proportion of the data could be retained.

After automatic artefact rejection according to the above criteria, the data was always manually reviewed to check if bad segments were correctly excluded. Bad channels were then replaced with data using spherical spline interpolation, which estimates the electrical potentials for the deleted channels from the remaining electrode sites. After the rejection of trials with artefacts, no participant had less than 20 trials per condition. The average rejection rate was 37.5% per condition (*SD* = 2.58). The remaining trials were averaged in each condition for each subject.

7.5 Statistical analysis procedures

The time windows were defined on the basis of the literature and visual inspection of the averaged data. Mean amplitude within a given time window was used as the dependent measure in the statistical analysis.

- The time windows for testing the effects after the onset of the word following the embedded verb were: 100-250 ms (ELAN), 300-500 ms (LAN), 400–700 ms (N400 1), 500-800 ms (N400, 2), 700-1000 ms (P600, 1), 850-1000 ms (P600, 2)
- The time windows for testing the effects after the onset of the word *that* were: 100-250 ms, 200-400 ms, 400–600 ms, 600-1000 ms (SAN)
- The time windows for testing the effects after the onset of the embedded verb were: 100-250 ms, 200-400 ms, 400-650 ms (negativities), 650-1000 ms (P600)

Nine groups of electrodes consisting of adjacent channels were selected for statistical analysis (Figure 1 shows the layout of the 128-electrode Geodesic Sensor Net). The clusters were defined in a way that enabled the exploration of distributional differences in the left-right and in the anterior-posterior dimensions: the two spatial factors laterality (left – mid - right) and anteriority (anterior – central – posterior) were completely crossed to provide the basis for the establishment of the nine electrode groups. The channels for each group were selected from among all the electrodes in the relevant regions of the scalp as defined by the two spatial factors. Eye electrodes and electrodes in marginal locations were not considered. Each group consisted of 6 electrodes except for the Mid Posterior cluster that had 5 channels. Mean amplitude for each cluster was computed by averaging the channels included in each cluster.

The following electrodes made up the nine clusters:

- Left Anterior: 23, 24, 26, 27, 28, 33
- Left Central: 35, 36, 40, 41, 45, 46
- Left Posterior: 53, 54, 59, 60, 61, 67
- Mid Anterior: 4, 5, 11, 12, 16, 19
- Mid Central: 7, 31, 55, 80, vref, 106
- Mid Posterior: 62, 71, 72, 75, 76
- Right Anterior: 2, 3, 118, 122, 123, 124
- Right Central: 102, 103, 104, 108, 109, 110
- Right Posterior: 77, 78, 79, 85, 86, 91

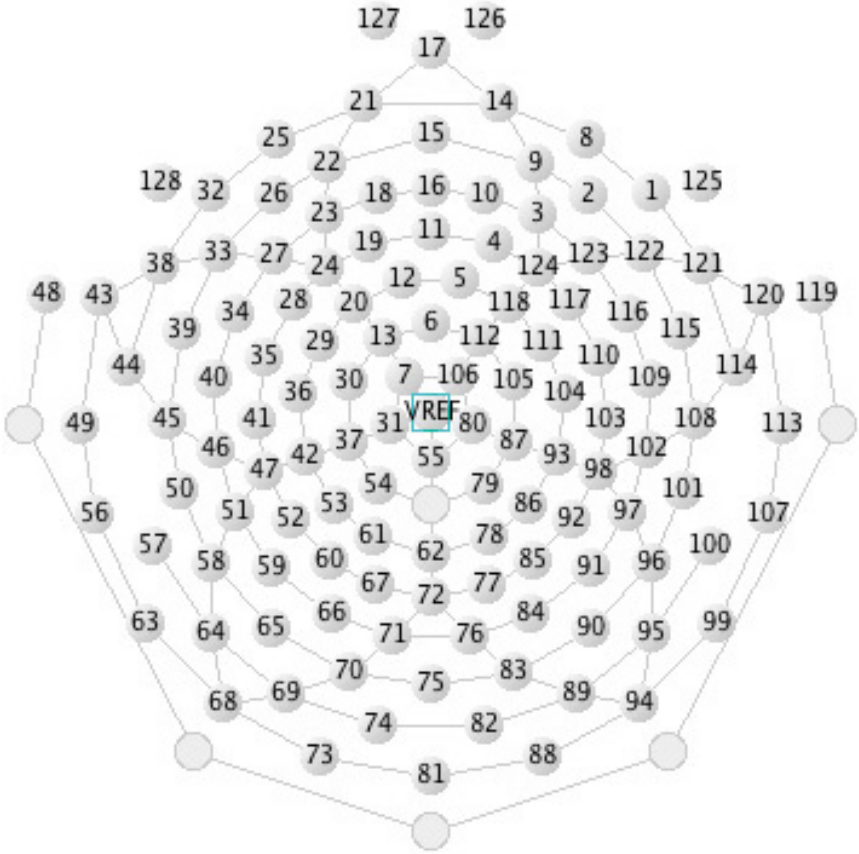

Fig. 1. Layout of the 128-electrode Geodesic Sensor Net.

In line with the formulation of separate experimental hypotheses for specific contrasts in Sections 6.3.1, 6.3.2 and 6.3.3, statistical analyses were performed separately for the following comparisons. Abbreviations are used to denote the specific critical words that were involved in the comparisons:

1. comparisons were made between the words directly following the embedded verb in the filled gap condition (the extra NP: ENP for short), the grammatical gap condition (the prepositional phrase: abbreviated as PP) and in the grammatical object condition (the grammatically and semantically correct NP: abbreviated as CorrNP)

2. between the NP directly following the embedded verb in the grammatical object condition (CorrNP) and the equivalent NP in the semantic anomaly condition (SemNP)

3. between the NP directly following the embedded verb in the correct subject-relative condition (CorrNP2) and the equivalent NP in the semantically anomalous subject-relative condition (SemNP2)

4. between the words *that* in object-relative *that* (ObjT), complementizer *that* (CompT) and subject-relative *that* (SubjT) conditions

5. between the embedded verbs in object-relative verb (ObjV), complement clause verb (CompV) and the subject-relative verb (SubjV) conditions

Repeated-measures ANOVAs were computed with the factor Condition and the two regional factors Laterality (levels: Left, Right and Mid) and Anteriority (levels: Anterior, Central and Posterior) in each time window. The distribution of Laterality and Anteriority corresponded to the 9 electrode clusters. ANOVAs were performed separately for each of the 5 comparisons described above, and the levels of the factor Condition always corresponded to the relevant contrasts listed for the different types of comparisons (e.g. for comparison (1) the three levels were ENP, PP and CorrNP).

In case of at least marginal ($p < .1$) interactions between Condition and one of the regional factors, further ANOVAs were computed at each level of the relevant regional factor. If there was an at least marginal Condition × Laterality × Anteriority interaction, one-way ANOVAs were performed on each electrode cluster. If a significant effect was obtained at a specific level of a regional factor or at a specific cluster, pairwise comparisons were made between the levels of Condition in a post hoc Bonferroni corrected t-test. To control for the violations of sphericity assumption, the Greenhouse-Geisser correction was applied with more than one degree of freedom in the denumerator. In these cases the corrected probability level is reported together with the original degrees of freedom.

8. Results

8.1 Comprehension questions

On average, the participants answered 89% of the questions correctly (minimum: 80%, maximum: 93%), which indicates that they paid attention to and understood the stimulus sentences (The native speakers in Hestvik et al.'s (2007) study had an average accuracy rate of 71% on the experimental trials, although they had a shorter time to respond: 5 seconds as

opposed to 20 seconds in the present study). In order to check whether the participants were able to deal with the relative clause structure, accuracy rate was calculated separately for those questions that specifically required the interpretation of "who did what to whom" in the sentences. These 33 items out of a total of 60 questions received correct answers in 85 % of the time, suggesting that the subjects understood the semantic relationships in the relative clauses.

8.2 ERP data: comparisons involving the word following the embedded verb

Grand average ERPs to the critical words in the different conditions are shown for each comparison at selected electrode sites in Figure 2, 5, 6, 7 and 8. Each electrode cluster is represented by one channel, and the waveforms were lowpass filtered at 15 Hz for display purposes only. The discussion of the ERP data for each comparison is always introduced by observations that were made on the basis of visual inspection of the grand average topoplot including all the 128 recording sites.

8.2.1 Filled gap, grammatical gap and grammatical object conditions

Visual inspection suggests an enhanced positivity for ENP (the extra NP in filled gap) compared to PP (the prepositional phrase in grammatical gap) and CorrNP (the NP in grammatical object), i.e. the voltage is positive going in the ERPs for ENP relative to the control conditions, starting around 800-900 ms (see Figure 2). This effect is visible over the posterior and central regions of both hemispheres. The positivity is more pronounced on left hemisphere electrodes; however, this difference was not confirmed by the statistical analysis. ENP does not show any reliable negativity effect throughout the whole epoch suggesting that neither an ELAN nor an N400 has been obtained.

The examination of the ERP waveforms of individual subjects revealed considerable within-group variation in these conditions. In order to investigate the nature of this variation, the results of those subjects whose electrophysiological responses showed a similar tendency were averaged together. In this way, three different learner groups emerged concerning their tendencies toward the ERP components elicited by ENP relative to PP and CorrNP. Three subjects displayed an anterior negativity most prominent over the left hemisphere around 100-350 ms, which suggests that the filled gap yielded an ELAN in these individuals. This negativity was followed by a late posterior positivity. Figure 3 shows the average ERP for these three participants at channel 27 from the left anterior region; Hestvik et al. (2007)

obtained a similar early negativity at the corresponding electrode in their study (channel AF7, see Hestvik et al., p. 313). Five subjects showed a tendency for an N400: an enhanced negativity starting around 350-500 ms and lasting until 700-800 ms. The effect had a posterior distribution extending to both hemispheres and was largest in centro-parietal locations (see Figure 4). No P600 was visible following the N400. Finally, the remaining 6 subjects displayed a P600 without any negativity effect preceding it. The fact that only a P600 effect was visible in the grand average is not surprising since it was the component that most of the subjects displayed. As it was not possible to statistically test the above-described within-group variation due to the low number of cases in each group, only statistical results obtained with all the 14 subjects will be presented.

Repeated-measures ANOVAs comparing ENP, PP and CorrNP resulted in statistically significant differences only in the 850-1000 ms time range. In this latency window, there was a marginal Condition × Anteriority interaction ($F(4, 52) = 2.47$, $p =. 093$). Subsequent region analyses revealed a main effect of condition in posterior ($F(2,82) = 11.95$, $p < .001$) and central locations ($F(2,82) = 4.32$, $p < .02$). Post hoc comparisons showed that the condition effect was due to increased positivity over posterior sites for ENP in comparison with both PP ($p < .01$) and CorrNP ($p < .01$), and due to a similar positivity effect for ENP in central regions when compared to PP (marginal: $p = .05$) and CorrNP ($p < .05$).

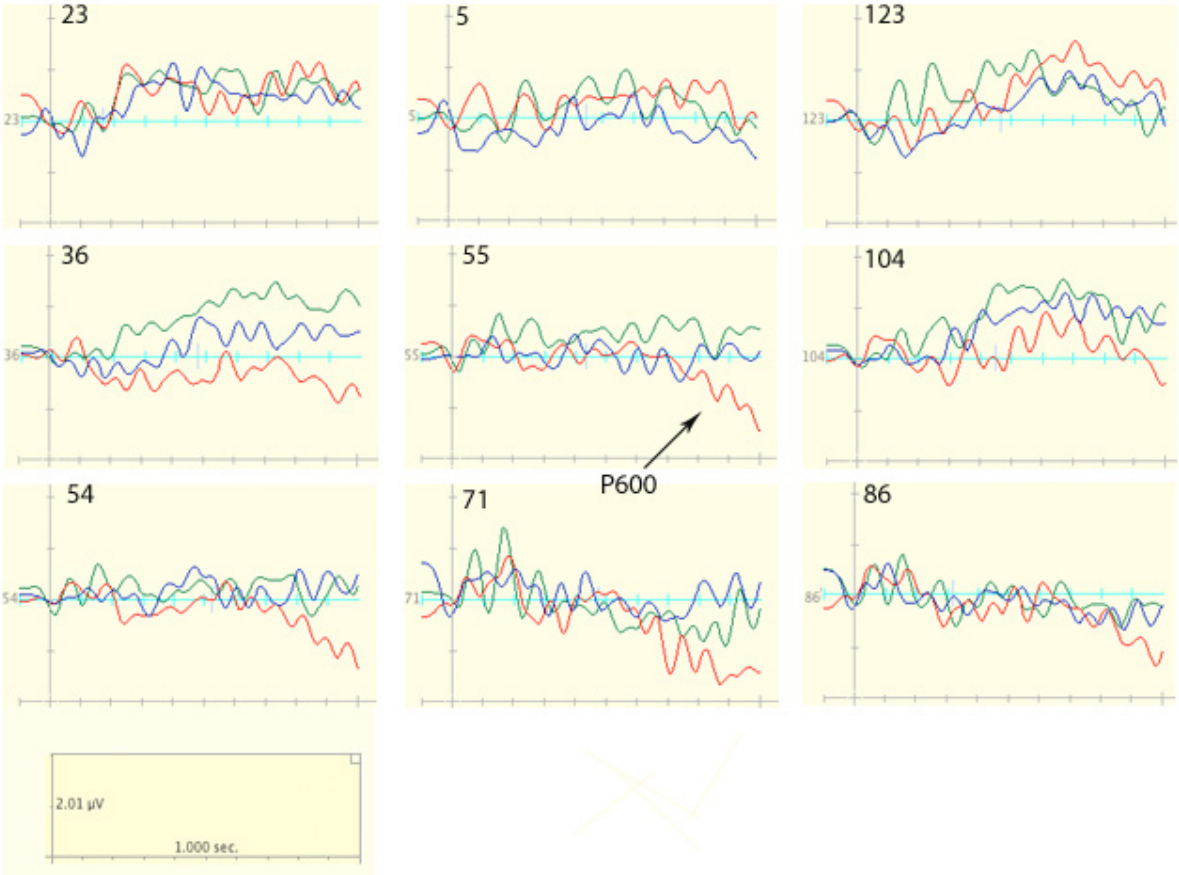

Fig. 2. Grand average ERPs to the word directly following the embedded verb at channels selected from each electrode cluster. Negative voltage is plotted upwards. The red line depicts the filled gap condition (ENP), the blue line the grammatical gap condition (PP) and the green line the grammatical object condition (CorrNP). There is a positive effect for the filled gap after 850 ms over central and posterior electrodes (P600).

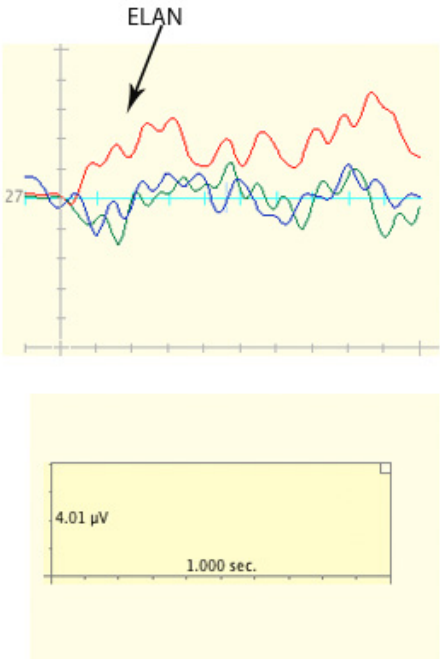

Fig. 3. Averaged ERPs of 3 subjects (the ELAN group) to the word directly following the embedded verb at channel 27 from the Left Anterior cluster. Negative voltage is plotted upwards. The red line depicts the filled gap condition (ENP), the blue line the grammatical gap condition (PP) and the green line the grammatical object condition (CorrNP). There is an early negativity for the filled gap between 80-350 ms (ELAN).

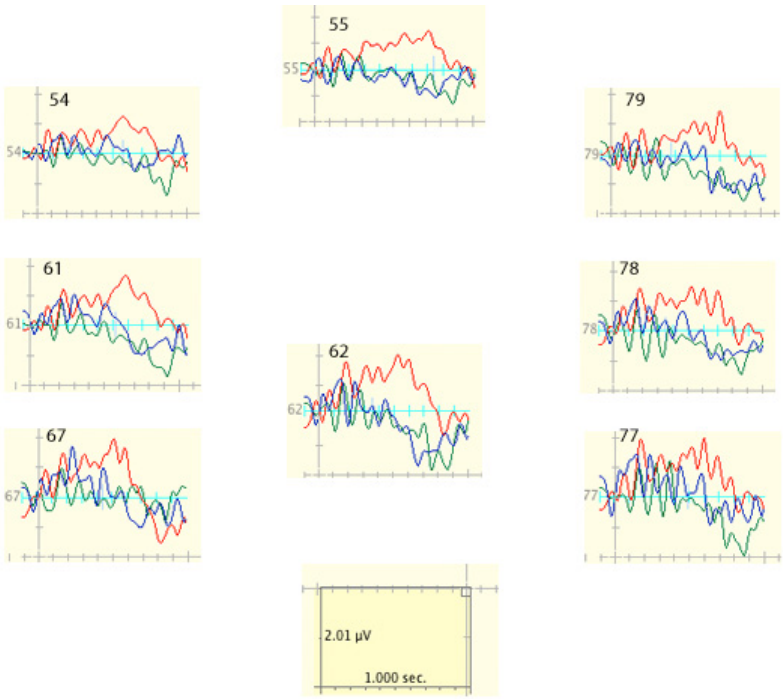

Fig. 4. Averaged ERPs of 5 subjects (the N400 group) to the word directly following the embedded verb at centro-parietal channels. Negative voltage is plotted upwards. The red line depicts the filled gap condition (ENP), the blue line the grammatical gap condition (PP) and the green line the grammatical object condition (CorrNP). There is a negativity for the filled gap starting around 350-500 ms and lasting until 700-800 ms (N400).

8.2.2 Grammatical object and semantic anomaly conditions

A positive waveform is observable for SemNP (the NP in semantic anomaly) relative to CorrNP (the NP in grammatical object) starting around 700 ms at most sites (see Figure 5). The positivity effect is more pronounced over the left hemisphere, and it is largest in the left central region. It also extends to central midline sites and to some channels in the central and posterior regions of the right hemisphere, but results of the statistical analysis did not confirm this observation. Unexpectedly, no negativity effect for SemNP is apparent in the N400 time window.

The statistical analyses revealed significant differences between SemNP and CorrNP in the 700-1000 ms time period. As far as the earlier time ranges are concerned, no reliable difference could be verified between the two sentence types. In the 700 - 1000 ms latency range, a marginal Condition × Laterality interaction was present ($F(2, 26) = 2.63$, $p = .092$), which was due to a positivity in the SemNP condition. This positive effect was significant only over the left hemisphere ($F(1, 41) = 6.15$, $p < .02$).

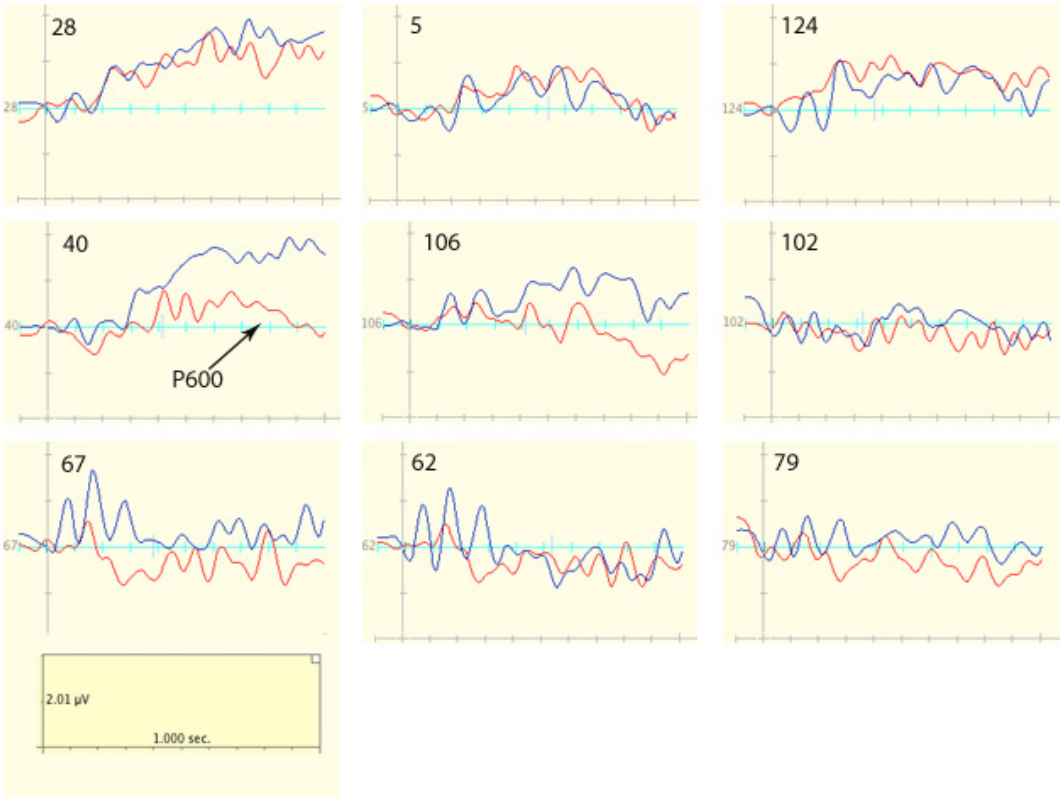

Fig. 5. Grand average ERPs to the NP following the embedded verb in the grammatical object and the semantic anomaly conditions at channels selected from each electrode cluster. Negative voltage is plotted upwards. The red line depicts the semantic anomaly condition (SemNP) and the blue line the grammatical object condition (CorrNP). There is an enhanced positivity for semantically anomalous NPs after 700 ms (P600).

8.2.3 Correct subject-relative and semantically anomalous subject-relative

Visual inspection suggests that SemNP2 (the NP in semantically anomalous subject-relative) elicited a larger negativity than CorrNP2 (the NP in correct subject-relative) after 700 ms at most electrodes (see Figure 6). This effect has a wide distribution involving several posterior and central electrodes, and it even extends to anterior sites, especially over the right hemisphere. The negativity is characterized by a clear centro-parietal maximum, and the effect starts already around 500 – 600 ms in this region. Nevertheless, such topographic differences were not supported by the statistical analysis: between 700-1000 ms, SemNP2 was significantly more negative than CorrNP2 ($F(1,13) = 4.89$, $p < .05$). There was no interaction of Condition with any of the regional factors.

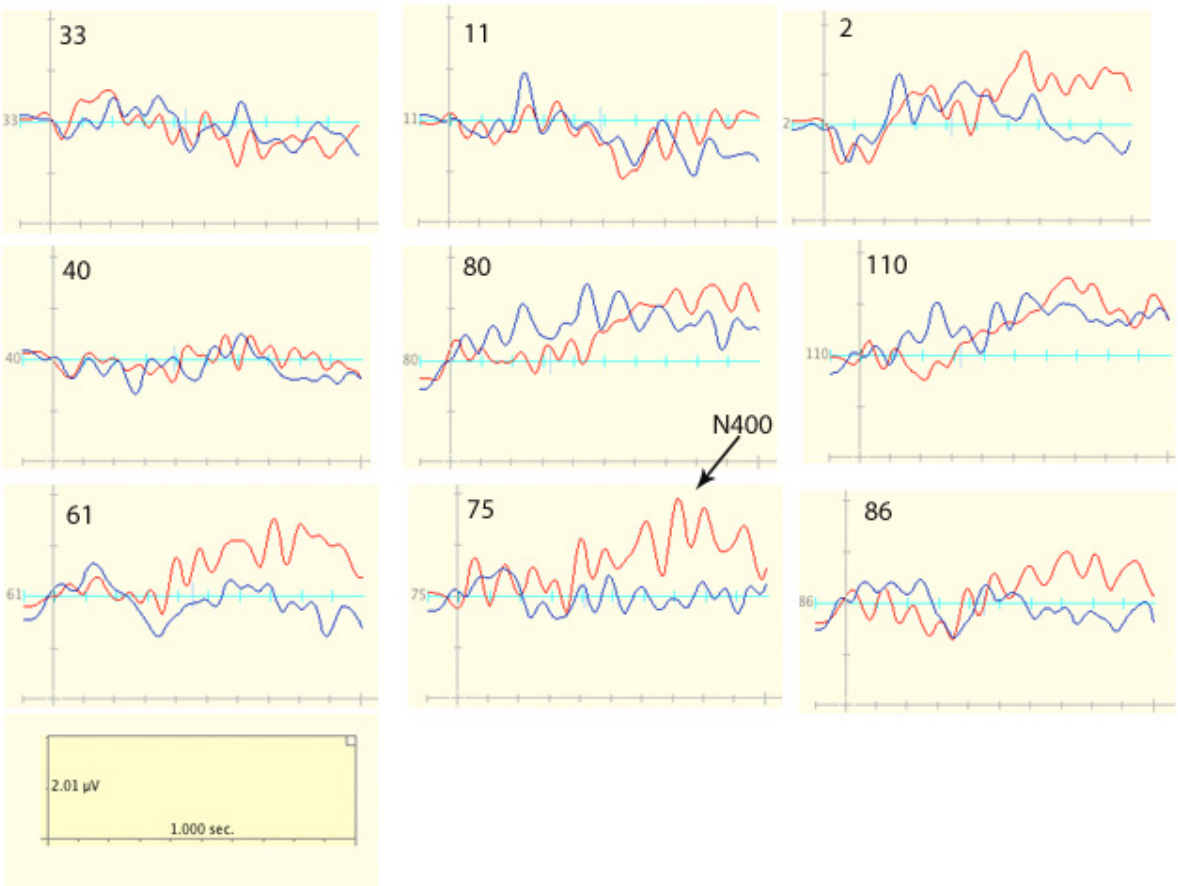

Fig. 6. Grand average ERPs to the NP following the embedded verb in the correct subject-relative and the semantically anomalous subject-relative conditions at channels selected from each electrode cluster. Negative voltage is plotted upwards. The red line depicts the semantically anomalous subject-relative condition (SemNP2) and the blue line the correct subject-relative condition (CorrNP2). There is a broadly distributed negative effect for semantically anomalous NPs (N400).

8.3 ERP data: comparisons involving the word *that*

Visual inspection indicates that the waveforms elicited by ObjT (object-relative *that*) and SubjT (subject-relative *that*) are very similar during the first 300 ms (see Figure 7). After 400 ms, ObjT shows a negative effect relative to both CompT (complementizer *that*) and SubjT. This effect continues until the end of the epoch and it has a broad frontal distribution. In posterior locations, SubjT displays a greater negativity compared to the two other conditions. The posterior negativity also has an onset around 400 ms, but it has a shorter latency of about 200 ms.

Repeated-measures ANOVAs indicated no difference between ObjT, CompT and SubjT in the two early time windows. Between 400–600 ms, a Condition × Anteriority interaction was found ($F(4, 52) = 4.92$, $p < .02$). Subsequent tests revealed a Condition effect for Anterior ($F(2, 82) = 6.50$, $p < .01$), Central ($F(2, 82) = 4.82$, $p < .05$) and Posterior (F(2,

82) = 11.78, $p < .001$) regions. Anteriorly, the condition effect was due to an increased negativity for ObjT as compared to CompT ($p < .02$) and SubjT ($p < .01$). The same tendency was found for central regions: ObjT was more negative than both CompT ($p < .05$) and SubjT ($p < .01$). Pairwise comparisons did not reveal any statistically significant differences between CompT and SubjT at the anterior and central regions ($p > .1$). In contrast, the condition effect in posterior locations was found to reflect the fact that SubjT was more negative than both ObjT ($p < .01$) and CompT ($p < .01$), whereas no difference between SubjT and CompT was found in this region.

Between 600-1000, there was a Condition × Anteriority interaction ($F(4, 52) = 3.52$, $p < .05$). Follow-up ANOVAs showed that the condition effect was present over anterior ($F(2, 82) = 7.99$, $p < .002$) and central ($F(2, 82) = 5.65$, $p < .01$) sites. Pairwise comparisons at the anterior region revealed a negativity for ObjT compared to both CompT ($p < .01$) and SubjT ($p < .05$). Centrally, ObjT was more negative than CompT ($p < .05$) and SubjT ($p < .01$), whereas CompT and SubjT did not differ in this region.

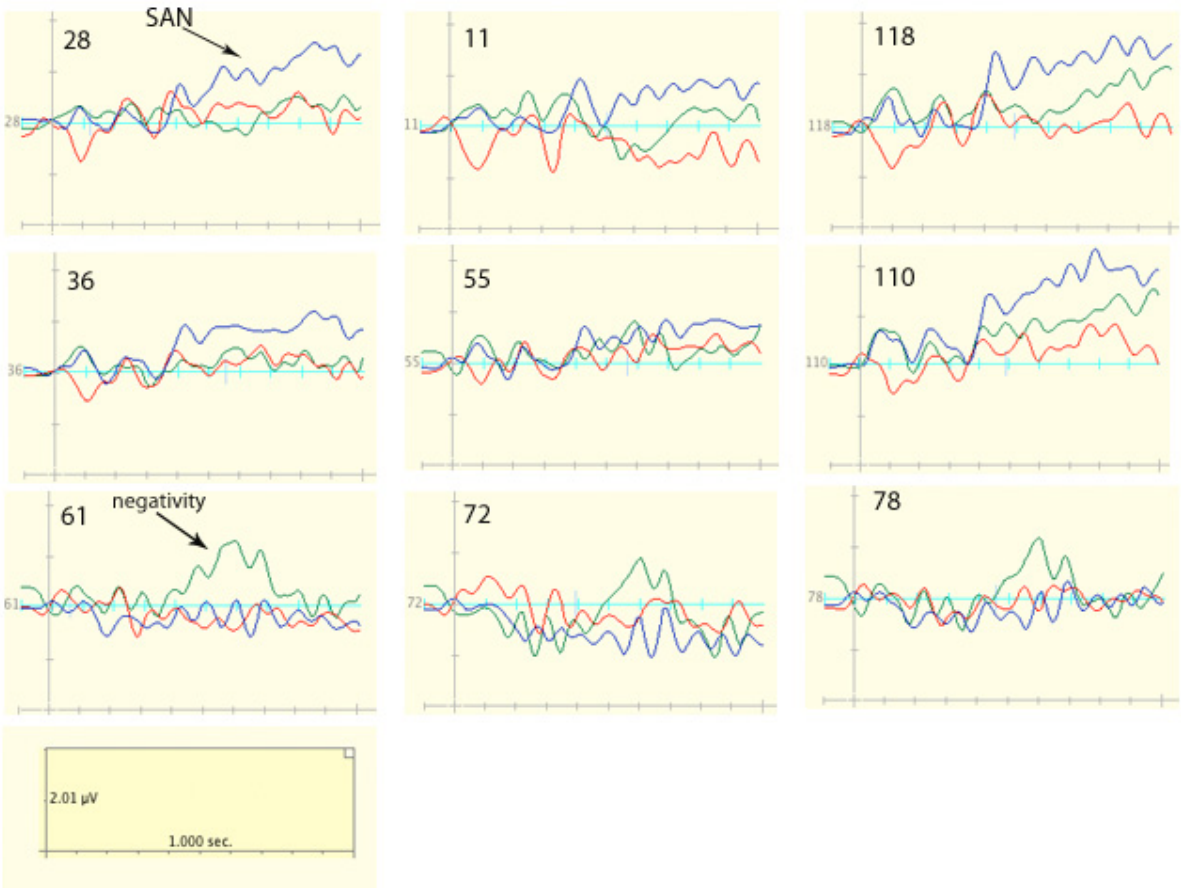

Fig. 7. Grand average ERPs to the word *that* at channels selected from each electrode cluster. Negative voltage is plotted upwards. The red line depicts the complementizer *that* condition (CompT), the blue line the object-relative *that* condition (ObjT) and the green line the subject-relative *that* condition (SubjT). There is an enhanced negativity for the object-relative *that* after 400 ms over anterior and central channels (SAN), and a posterior negativity between 400-600 ms for the subject-relative *that* is probably attributable to lexical differences in the following words.

8.4 ERP data: comparisons involving the embedded verb

ObjV (object-relative verb) shows hardly any differences in comparison to CompV (complement clause verb) throughout the whole epoch (see Figure 8). At the same time, a larger negativity is observable for SubjV (subject-relative verb) sentences relative to the other two conditions over the entire right hemisphere, especially over central and posterior sites. The negativity has an onset around 400-600 ms, and it continues until the end of the epoch.

Repeated-measures ANOVAs found significant effects involving the factor Condition in the 400-650 ms and the 650-1000 ms time period. In detail, there was a marginal Condition × Laterality × Anteriority interaction between 400 and 650 ms ($F(8, 104) = 2.20$, $p = .086$. Resolving the three-way interaction revealed that the condition effect was restricted to

the Right Central cluster ($F(2, 26) = 9.28$, $p < .002$). Here, SubjV was more negative than both ObjV ($p < .01$) and CompV ($p < .01$), whereas SubjV and CompV did not differ from one another ($p > .1$). Between 650-1000 ms, a Condition × Laterality interaction was present ($F(4, 52) = 3.00$, $p = .05$). Subsequent hemisphere analyses found a condition effect only over the right hemisphere ($F(2,82) = 8.18$, $p < .002$), where SubjV displayed a negativity compared to both ObjV ($p < .01$) and CompV ($p < .05$). ObjV and CompV did not differ in this time range either ($p > .1$).

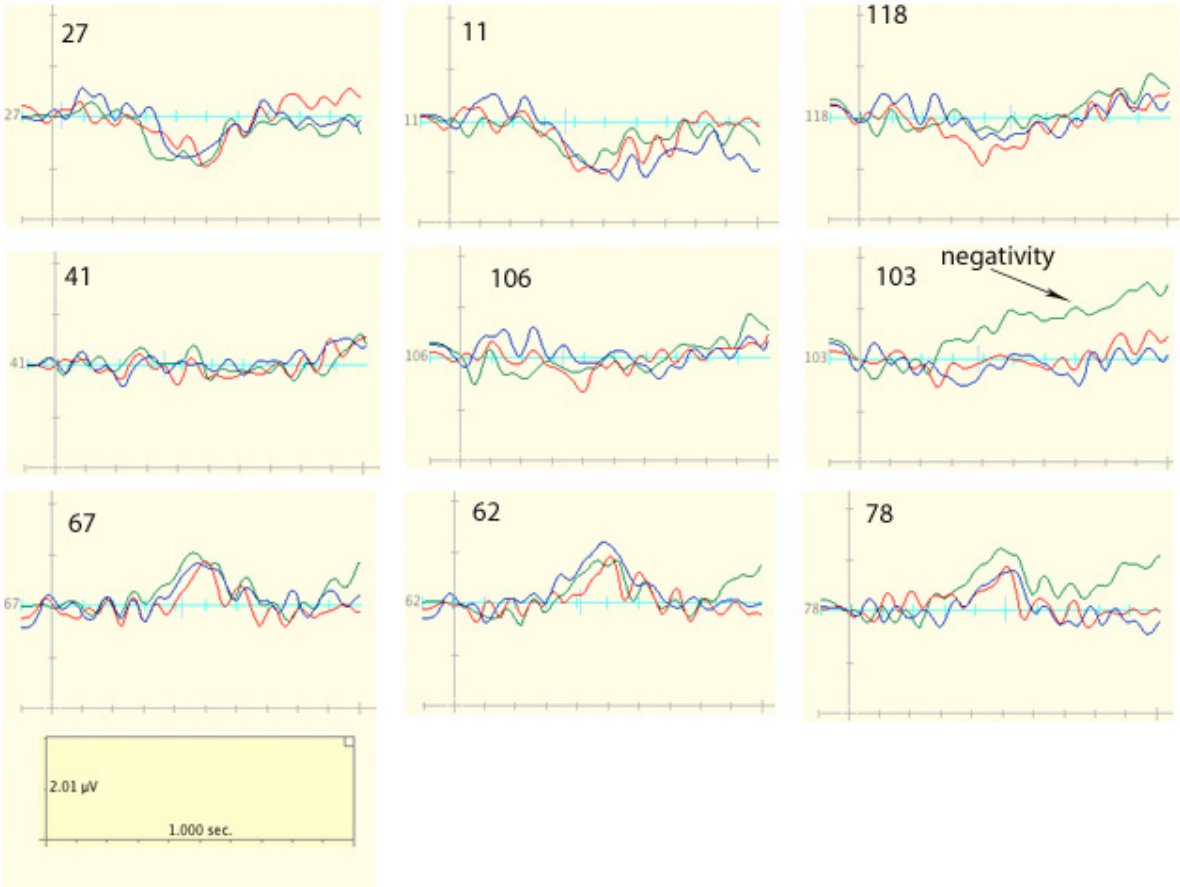

Fig. 8. Grand average ERPs to the embedded verb at channels selected from each electrode cluster. Negative voltage is plotted upwards. The red line depicts the complement clause verb condition, the blue line the object-relative verb condition and the green line the subject-relative verb condition. The subject-relative verb is more negative after 400 ms over right central electrodes, and over other right hemisphere channels as well after 650 ms.

9. Discussion

9.1 Processing the filled gap

Relative to the grammatical gap and grammatical object sentences, the filled-gap condition elicited a late positivity between 850–1000 ms with a central-posterior distribution. The latency and distribution of this positivity suggest that it is a somewhat delayed P600. The participants, therefore, differed from the NSs of Hestvik et al.'s (2007) study in that the ungrammatical NP did not yield an ELAN. At the same time, no N400 was observed either as opposed to the predictions of the SSH.

The absence of the N400 suggests that the participants did not experience problems with semantic integration at the filled gap position, which is possible if we assume that the learners initially identified the extra NP as the direct object of the preceding verb: since the additional NP was always semantically compatible with the verb, thematic interpretation did not cause any processing difficulties. Locally, the word order also supported this interpretation, as an overt direct object would have appeared after the verb in a main clause or non-relativized subclause. Section 4.1 reviewed studies that provided evidence for L2 learners' use of the active filler strategy to associate the filler with an appropriate subcategorizer as soon as possible. Nevertheless, the unproblematic integration of the additional NP at the filled gap position indicates that the participants of this study had not yet made an association between the clause-initial object NP and the verb before the extra object was reached. This delay might be explained by assuming that the processing of the verb took longer for the learners than for the NSs in Hestvik et al.'s (2007) study. In addition to compositional integration between the preceding subject and the verb, processing the embedded verb in the object-relative clauses of the present experiment presumably involves evaluation of the requirements of the verb concerning further arguments and the retrieval and/or reactivation of the clause-initial NP in working memory to associate it with the verb as its object. The L2 learners might have been able to accomplish only a subset of these processes before they encountered the extra argument; specifically, there was no evidence for the integration of any of the properties of the clause-initial object NP before that point. ERP responses obtained at the verb position support this assumption: no difference was found between the object-relative verb and the complement clause verb despite the fact that by the time the object-relative verb was reached, both verbal arguments had already occurred in the sentence, whereas the object appeared only after the verb in the complement clause condition.

It is, nevertheless, unlikely that the clause-initial NP simply remained unintegrated. First of all, responses to the comprehension questions indicated that the participants correctly understood the semantic relationships in the grammatical object-relative clauses including the role played by the clause-initial NP. It seems that even if with some delay, the clause-initial NP was integrated as the object of the embedded verb in correct sentences. It is, unfortunately, not possible to check if the participants interpreted the extra NP as the object of the verb in the ungrammatical sentences as comprehension questions intentionally avoided this region of the clause. Second, the filled gap elicited a P600 component associated with syntactic reanalysis and repair processes, which possibly reflected learners' attempt at integrating the clause-initial NP after it has been reactivated in or retrieved from memory with some delay. Since no N400 was found, which would have indicated problems with semantic interpretation; the reanalysis process was probably triggered by the impossibility of integrating this NP into the previously established syntactic structure as a second NP after the verb. As it was described in Section 5.1, two NPs directly following the verb in this context creates a phrase structure which is not well-formed, and the P600 has been previously observed with phrase structure violations (e.g. Friederici et al., 1996). It is, however, not possible to decide if the learners postulated a structurally present gap after the verb to which they tried to link the clause-initial NP or the integration of the NP into the structure was not mediated by such gaps. Alternatively, the P600 indicated a syntactic reanalysis process that took place even before the clause-initial NP was reactivated in or retrieved from memory: the extra NP occupying the canonical position of the clause-initial NP was first interpreted as the direct object of the verb based on semantic congruity. It was then evaluated if the syntactic structure of the sentence supported this interpretation as well, and the P600 was elicited by the recognition that an overt object NP in the present relative clause structure cannot possibly appear after a transitive verb taking only one object NP.

The above-outlined interpretation of the results is partly consistent with the shallow structure hypothesis; the learners initially seem to have processed the semantically congruous but ungrammatical NP as the object of the verb, which would imply the use of a predominantly semantic strategy to associate the verb with its arguments. However, the learners displayed sensitivity to syntactic information as well since their interpretation of the extra NP as the object eventually led to repair or reanalysis processes concerning the syntactic structure. These findings do not support the assumption that learners rely less on syntactic information than NSs. Instead, the difference between native and L2 parsing mechanisms might be essentially quantitative in nature, and it might be related to the speed

with which language input is processed: slower processing at the verb position could have contributed to the non-native-like response pattern L2 learners displayed for the filled gap. From this perspective, the three learner subgroups identified in this study on the basis of their ERP responses to the filled gap (ELAN, N400, P600; see Section 8.2.1) represent different stages in a continuum defined by speed of sentence processing.

First, the ELAN group appear to have processed the filled gap sentences of the present experiment in a native-like manner since they displayed the same ERP response as the NSs in Hestvik et al.'s (2007) study. In the N400 group, however, no ELAN was obtained indicating that they were unable to identify the phrase structure violation early enough for this component to appear. Nevertheless, processing at the verb was faster than in the P600 group, and they managed to integrate the clause-initial NP with the verb before the interference of the extra NP. The N400 response probably reflected their subsequent attempt at the semantic integration of the additional argument. The P600 indexing difficulties with accommodating an extra object in the phrase structure representation was likely to be too small to be visible with only five subjects: for example, the P600 following an N400 elicited by superfluous arguments in Friederici and Frisch's (2000) study also had a reduced amplitude.

Segalowitz's (2003) discussion of the relationship between fast processing and automaticity in L2 acquisition might shed light on the way variations in speed of processing can lead to such apparently qualitative differences as the presence or absence of the ELAN. According to Segalowitz, language comprehension involves the operation of a wide range of mechanisms including serial, parallel and cascading processes. For some of these operations, practice can lead to extremely fast processing while certain mechanisms will always be relatively slow. As a result of training, the former kind of activities might achieve such a speed that other mechanisms can no longer have an influence on their operation, and it also leads to a change in the relative order of the processes (Segalowitz, 2003). From this perspective, the ELAN might reflect a mechanism involved in sentence comprehension that operates very rapidly in NSs as a result of practice, and which is independent from the influence of other types of information due to its speed. Friederici (2002) proposed that the ELAN correlates with an autonomous first stage of sentence parsing during which a phrase structure representation is built up on the basis of word category information. Adopting this account, it is suggested that the observed differences between the NSs and the ELAN group of this study versus the N400 and the P600 group are related to the speed of phrase structure building operations during real-time sentence comprehension. In case of a language acquired

later in life and not used extensively on an everyday basis, these processes might not have been practiced to the degree that is necessary for achieving a significant increase in speed relative to other component mechanisms underlying language comprehension. Consequently, these mental activities might take place simultaneously with or even later than other operations such as semantic integration processes. It might thus appear that learners such as those constituting the N400 and P600 group in the present experiment rely more on semantic information during parsing, while they in fact also make use of the same syntactic cues as NSs do.

Similarly to the findings of this study, Felser et al. (2012) observed that the influence of semantic information on non-native sentence processing appears to precede the application of syntactic cues. However, their explanation did not make reference to differences in speed of processing as they suggested that some part of learners' grammatical knowledge might be represented in a form that is not available for earlier parsing operations resulting in the use of shallow parsing mechanisms (see Section 4.1). It is equally plausible, however, that there is no difference between learners and NSs in terms of their linguistic knowledge representations, but structure building processes that make use of certain part of this knowledge take place relatively late during the comprehension of a less-practiced language. The finding of the present study that some of the late learners displayed ERP responses indicative of the rapid operation of structure building mechanisms demonstrates that practice can indeed lead to native-like performance concerning the relevant processes.

9.2 Semantic anomaly

As discussed in Section 6.2, sentences involving clear semantic anomalies were included among the experimental items to be able to assess whether these violation types elicit the expected N400 in the participants of the present study. The semantically incongruent NP in the semantically anomalous subject-relative sentences indeed yielded a broad bilateral negativity between 700 and 1000 ms relative to the correct subject-relative condition. Several studies with NSs using either visual or auditory stimuli have found effects that were taken to be instances of the N400 despite their broad distribution extending not only to central-posterior sites but to anterior regions as well (e.g.: Friederici & Frisch, 2000; Frisch et al., 2004; Hahne, 2001). The latency of the present negativity is delayed even in comparison with the N400 effects that have been previously reported with L2 learners. Part of this additional delay is most likely the result of the fact that the ERPs were time-locked to the determiner

preceding the noun and not to the noun itself. Consequently, the obtained component appears to be a delayed and widely distributed instance of the N400, and the observation of this effect in semantically anomalous sentences provides support for the interpretation of the absence of the N400 in the filled gap condition as indicating the fact that learners did not experience problems with semantic integration at this point of the sentence.

Surprisingly, the same kind of violation constituted by the same lexical items did not lead to an N400 in the semantic anomaly condition relative to the grammatical object items. In this case, a left lateralized late positivity was obtained between 700 – 1000 ms. The component appeared in the slightly delayed latency window of the P600 as expected with L2 learners, while it had a different topography than the typical posterior P600 effect elicited by grammatical violations. The present positivity had a wider and more frontal distribution, which is more consistent with the topography of those late positivites that have been observed for the processing of grammatically correct structures at the point where an already established phrase structure representation was presumably revised (see section 3.1). Furthermore, while the statistical analysis indicated that the positivity was restricted to the left hemisphere, visual inspection suggested that it was present over several midline and right hemisphere channels as well, supporting the assumption that the observed effect is a P600.

A P600 component has been reported for certain types of grammatically correct sentences involving obvious semantic violations (see Section 3.1). However, the present effect does not seem to be a typical case of the semantic P600: there was no close semantic fit between the noun and the verb, and the experimental items did not include animacy reversals either as the anomalous noun would not have been a plausible argument of the verb with opposite thematic role assignments. For example, both the sequence *the painter scared the document* and *the document scared the painter* are semantically odd. Nevertheless, some previous studies demonstrated that close semantic relationship between the verb and the arguments is not a crucial factor for eliciting the semantic P600 (Hoeks et al., 2004; Kuperberg et al., 2007; Stroud & Phillips, 2011). According to Kuperberg et al. (2007) and Stroud and Phillips (2011), the obtained late positivity in their studies potentially reflected thematic processing difficulties deriving from the fact that the noun (*eggs* in (38)) was not simply an implausible argument of the verb (*plant*), but it also violated its specific requirement of an animate subject.

(38) For breakfast the eggs would plant flowers in the garden. (Kuperberg et al., 2007, p. 226)

To account for the absence of the N400 in these cases, Kuperberg et al. suggested that "semantic integration … may have been attenuated by the cost in processing reflected by the P600" (p. 235), i.e. the detection of the animacy-requirement violation led to some reanalysis processes concerning the establishment of alternative thematic relations, which also involved an attempt at reanalysing the syntactic structure to be consistent with the alternative thematic relationships. Discussing the factors that can potentially provoke such reanalysis processes, Kuperberg et al. argued that certain characteristics of the stimulus material (or the related task) might suggest that it is possible to make sense of the anomalous sentences. For example, the structure of the surrounding stimulus sentences, presence of some context preceding the violation or the use of plausibility judgement task might motivate participants to asses alternative thematic relationships.

With regard to the present study, the semantic anomalous item was indeed an inanimate noun in each case, but it did not always constitute a general violation of the animacy requirements of the verb. Only about one-third of the relevant verbs might have been regarded as specifically requiring an animate object (e.g. *thanked, scared),* a few seemed to be more plausible with animate objects than with inanimate ones (e.g. *hugged*) while the rest was apparently as likely to appear with inanimate objects as with animate ones (e.g. *watched, tapped*). Nevertheless, sentences in the grammatical object, correct subject-relative and filled gap conditions always involved a noun denoting a profession after the embedded verb. This property of the stimulus material could have generated a strong expectation of an animate, moreover human object in the postverbal position of the semantically anomalous sentences, which might have provoked attempts at reanalysis even with verbs that would have been completely acceptable with inanimate objects in different contexts. In addition, the grammatical object, grammatical gap and correct subject-relative conditions involved both grammatically and semantically correct sentences with highly similar structure to the semantically anomalous items, which might have suggested that it is possible to make sense of the semantically anomalous sentences as well.

These factors, however, do not explain why the object NP in the semantically anomalous subject-relative items elicited the predicted N400 instead of a P600. An important difference between the two conditions is that in the semantic anomaly sentences, there is always a clause-initial NP (a profession noun) acting as the subject of the main clause verb. This element would be an appropriate object of the embedded verb in terms of general semantic and pragmatic fit and also considering the other experimental sentences in which

the embedded verb takes an object signifying a profession. Based on the observation that the presence of a previous NP providing a good semantic fit seemed to be a decisive factor for eliciting the P600 in the semantic anomaly condition, it is tentatively suggested that the profession noun encountered earlier during sentence processing motivated the listener to recheck the already constructed interpretation and evaluate if it is possible to establish an argument-predicate relationship between the initial NP and the embedded verb. P600 might have reflected the participants' attempt at reanalysing the already constructed phrase structure representation to reconcile this alternative interpretation with the relationships dictated by the syntactic structure of the sentence.

The proposed explanation for the P600 effect is consistent with the SSH in that the L2 learners apparently disregarded the interpretation unambiguously suggested by the syntax when they considered the clause-initial NP as a potential argument of the embedded verb on the basis of semantic fit. The observed P600, however, indicate that they still recognized the incompatibility of this relationship with the actual syntactic structure: it appears that the application of semantic information might initially precede the consideration of syntactic cues, but both types of information play a role in constructing the sentence interpretation. In addition, the fact that the observed effect can be argued to parallel NSs' responses to specific kinds of violations yielding a semantic P600 does not support the assumption of qualitatively different sentence parsing mechanisms in native and L2 language processing.

9.3 Processing the word *that* in filler-gap dependencies

Although the sentences in the two relative-clause conditions were identical until the word following *that*, responses yielded by the subject-relative *that* and the object-relative *that* diverged after 400 ms. This difference was probably due to the fact that by that time, the participants have identified one of these elements as the word standing for the subject of an upcoming relative clause and the other as the word standing for the object of a relative clause. This assumption is supported by the observation that the word *that* has a very short duration: 135 ms on average in the present stimulus material. Moreover, in case of the experimental sentences of this study, relative *that* is always immediately followed by the definite article *the* when it introduces an object-relative clause instead of a subject-relative one. The shortness of the article probably enables the fast recognition of the fact that an overt relative-clause subject follows and, consequently, *that* must stand for some other argument coming later in

the sentence. The lack of a following definite article, in turn, instantly identifies *that* as the element standing for the relative-clause subject.

Between 400 and 600 ms, a bilateral anterior-central negativity was obtained for the object-relative *that* compared to the two other conditions. In the same time range, a negative effect was found also for the subject-relative *that*, but with a wide posterior distribution. The negativity for the object-relative *that* was present even after 600 ms, lasting until the end of the epoch, and it had a similar anterior-central distribution as in the earlier time period. The negativity elicited by the object-relative *that* may be connected to the sustained anterior negativity (see Section 5). The timing of the effect is consistent with this assumption: the SAN observed by Fiebach et al. (2002) also started 400 ms after word-onset. Nevertheless, one difference between the characteristics of the SAN effects observed previously (described in Section 5) and the present effect is that the component obtained in this study is not restricted to the left hemisphere. Fiebach et al. (2002) reported a similar broader distribution with NSs who had a low working memory span, and they interpreted the difference as reflecting the greater working memory resources this group of participants need to use for establishing the filler-gap dependency. In case of the L2 learners, greater working memory load might not have been created by the storage of the filler as such but it might have derived from the need to maintain this element in memory in addition to the generally higher processing demands of comprehending an L2.

A further difference between the present negativity and the SAN obtained in previous NS studies is that this component often appeared only on the element immediately following the relative pronoun (Fiebach et al. 2002; Kaan et al. 2000; Kluender & Kutas, 1993). In all these studies, the relative pronoun was a *wh*-question phrase (*who* in (39)), and as such, it did not have a specific referent mentioned earlier in the sentence.

(39) Have you forgotten who he dragged ____ to the movie that weekend? (Kluender & Kutas, 1993, p. 200)

By contrast, the relative pronoun *that* in the present study had a specific referent and it was the same as that of the noun occurring immediately before it. From this perspective, obtaining a SAN at the relative pronoun *that* is similar to observing this component at the next word after a *wh*-question phrase; the SAN starts in both cases when the absence of an appropriate subcategorizer and the presence of an element marking a relative clause make it clear that the

preceding word cannot be integrated immediately into the sentence representation and storage in working memory is required until its subcategorizer is encountered.

A negative effect was obtained also for the subject-relative *that*; nevertheless, the posterior distribution of this component challenges the assumption that it is related to the SAN. Phillips et al. (2005) also found a posterior negativity at the *wh*-element introducing the filler-gap dependency but attributed the effect to differences between the experimental conditions in terms of the word following the *wh*-phrase. Similar lexical differences were present in the current study as well: the subject-relative *that* was followed by an open-class word (a verb) while the item after the *that* in the other two conditions was always a closed-class word (a determiner). Since open-class words yield a larger N400 (Kutas & van Petten, 2006), the posterior negativity at the subject-relative *that* might have simply reflected an effect elicited by the next word. Consequently, and in accordance with the predictions for this condition outlined in the hypothesis section, no evidence was obtained for taking the filler into working memory at the subject-relative *that*, where the immediately following word is the embedded verb providing the filler with its interpretation.

9.4 Processing the embedded verb in filler-gap dependencies

In contrast to the expectations, no P600 effect was found for the object-relative verb relative to the other two conditions. As a matter of fact, the object-relative verb did not differ significantly from the complement clause verb in any time windows, which also makes it unlikely that the late positivity was simply attenuated by a sustained anterior negativity starting at the word *that* in the object-relative condition and continuing up until the verb. In this case, a negative effect would have been expected for the object-relative verb in the earlier latency windows.

Given the assumption that the amplitude of the P600 reflects syntactic integration costs as discussed in Section 5, the obtained results suggest that syntactic integrations taking place at the point of processing the object-relative verb were not more demanding than those occurring at the complement clause verb. This interpretation is consistent with the explanation that has been offered for the P600 effect elicited by the filled gap condition in Section 9.1. As learners' processing of the embedded verb is relatively slow, reactivation of the clause-initial object NP and its semantic and syntactic integration with the object-relative verb are delayed to an extent that no effect of processing complexity can be observed at this position relative to the complement clause verb: in both cases, it is only the preceding subject

NP that is integrated with the verb and operations involving the object NP take place only later.

The failure to obtain any significant difference between the object-relative verb and the complement clause verb in terms of ERP responses also implies that the anterior negativity observed at the word *that* in the object-relative clauses did not continue throughout the dependency. Given that the SAN-like effect at the object-relative *that* was taken to index storage of the filler in working memory, and it was argued that filler integration was delayed until the next word after the verb, a negativity would have been expected to be still present at the verb position. One possible explanation is that the length of the dependency was so short that the maintenance of the filler in itself after the relative pronoun did not induce a working memory load eliciting a significantly increased negativity: the amplitude of the SAN was indeed found to increase over the course of the dependency in NSs (Fiebach et al., 2002). Alternatively, the filler was taken into working memory at the word *that* but comprehending the rest of the sentence incurred such great processing costs for the learners that no resources could be allocated to active maintenance of the filler explaining the absence of the SAN. The evidence available does not enable us to distinguish between these contrasting accounts.

For the subject-relative verb condition, a right lateralized negativity effect was obtained in comparison with both the complement clause verb and the object-relative verb. The negativity started around 400 ms and lasted until the end of the epoch. With time, the distribution of the effect became wider: while it was restricted to the right-central electrodes between 400 and 650 ms, later it became significant over all right hemisphere electrode clusters. One possibility is that this effect is related to the SAN, which has been observed at the point where the dependency can be completed and was interpreted to index the retrieval of the filler from working memory (e.g. Felser et al, 2003; see Section 5). The present negativity and the SAN obtained in Felser et al.'s (2003) study indeed have the same onset time whereas they differ in their spatial distribution. Moreover, no SAN-like negativity was found previously at *that* in the subject-relative condition, which was interpreted to indicate that the filler was not taken into working memory due to the shortness of the dependency. Consequently, it appears that the present negativity and the SAN found at the subcategorizing verb with longer dependences reflect somewhat different cognitive processes. The subject-relative verb differs from both the complement clause verb and the object-relative verb in that the verb is not directly adjacent to its subject, suggesting that the negativity might reflect the influence of distance on the integration of the verb with its subject. This evidence does not

enable us to determine the exact nature of this influence: if it is related to syntactic and/or semantic integration or to working memory processes.

9.5 A potential confounding factor

Finally, it is necessary to discuss a factor that potentially could have complicated comparisons at the *that* and embedded verb positions: the sentence position effect, which entails that open-class words generally elicit gradually smaller N400s as the sentence progresses. The emerging sentence context presumably facilitates the processing of upcoming words that fit that context (Kutas & van Petten, 2006). With regard to the comparisons involving the embedded verb position, the subject-relative verb is the fourth word from the beginning of the sentence whereas the object-relative verb and the complement clause verb are the sixth and the seventh, respectively. It is possible, therefore, that a greater N400 related to the sentence position effect contributed to the observed negativity at the subject-relative verb; still, there is reason to believe that it cannot explain the entire effect. First, the long duration (400-1000 ms) and right lateralized distribution of the negativity elicited by the subject-relative verb are inconsistent with the characteristics of the N400s typically obtained with NSs. For example, the effect does not tend to be confined to one hemisphere, and the auditory N400 has even been reported to be larger over the left than the right hemisphere (Hagoort & Brown, 2000). Second, the negativity elicited by the subject-relative verb is also different from the N400 actually obtained for the semantically anomalous sentences of the present study: the latter effect was characterized by a wide bilateral distribution with a centro-parietal maximum as indicated by the visual inspection.

As far as the *that*-conditions are concerned, the complementizer *that* appears later in the sentence than the equivalent word functioning as a relative pronoun in the other two conditions (third versus fourth word). *That* is, however, a closed-class word suggesting that the sentence position effect reported with open-class items might not even play a role in these comparisons. Moreover, the observed negativity effect was present not only between the complementizer *that* versus the object-relative and subject-relative *that* but also between the object-relative *that* and subject-relative *that*, where the distance from the beginning of the sentence was identical.

10. Summary and conclusions

The aim of the present study was to investigate if L2 learners of English differ from NSs in their processing of filler-gap dependencies and if the potentially observed differences are consistent with the claims of Clahsen and Felser's (2006) *shallow structure hypothesis* according to which learners underuse syntactic information during sentence processing and establish a direct semantic association between the filler and the verb in filler-gap dependencies. To that end, ERPs were recorded at the word *that* and at the embedded verb in object-relative clauses, subject-relative clauses and in finite complement clauses. The effect of an extra NP incorrectly occupying the assumed gap position in object-relative clauses was also tested, and ERPs obtained in this filled gap condition were compared to the findings of a previous study with English NSs (Hestvik et al., 2007).

Results of the present study have the following implications for the strategies that L2 learners use to associate the filler with its subcategorizer during on-line filler-gap dependency processing. First, the absence of an ELAN response to the extra NP in the filled gap condition indicates that the learners, unlike the NSs in Hestvik et al.'s (2007) study, did not construct a structurally present gap in the canonical position of the filler as part of an initial, autonomous phrase structure building process. Furthermore, the failure to observe an N400 in this condition suggests that the learners did not immediately establish a semantic association between the verb and the clause-initial NP either. Instead, it was argued that learners processed the embedded verb slower than NSs: no difference was found between the object-relative clauses and finite complement clauses at the point of processing the embedded verb, indicating that the integration of the clause-initial NP with the verb was delayed until the following word in object-relative clauses. It was suggested that this delay in the filled gap sentences led to an initial association of the locally appropriate extra NP with the preceding verb and to a subsequent detection of the syntactic violation involved in this association as indicated by the obtained P600. Since ERP responses to the filled gap position apparently reflected the consequences of interpreting the extra NP as the object of the verb, the results of this study do not enable us to identify the exact mechanism that L2 learners use to link the filler to the subcategorizer in grammatical sentences.

Concerning working memory processes involved in the comprehension of filler-gap dependencies, a sustained anterior negativity was found at the word *that* in object-relative clauses but not in subject-relative clauses suggesting that the L2 learners took the filler into working memory at the relative pronoun when the immediately following word was not

expected to be an appropriate subcategorizer. The negativity apparently did not continue throughout the dependency until the integration of the filler, but it is not clear if this finding is indicative of differences between NSs and L2 learners concerning their abilities to allocate processing resources for the maintenance of the filler in working memory until the subcategorizer is reached. Further ERP studies should aim to clarify this issue.

With regard to the information types that learners make use of during sentence comprehension, the results suggest that certain semantic integration processes such as thematic role assignments might precede syntactic operations. At the same time, learners were apparently sensitive to syntactic information as well, and the findings are even consistent with the assumption that they constructed phrase structure representations: the P600 response for the filled gap was interpreted to reflect syntactic reanalysis or repair processes potentially provoked by the recognition that two NPs following the verb create a phrase structure which is not well-formed in the experimental sentences. Further evidence for learners' use of syntactic information was provided by the observation of a P600 effect without a preceding N400 for the semantically anomalous object NP in finite complement clauses: this finding was taken to indicate that the syntactic structure of the sentence imposed constraints on the reinterpretation of thematic relationships between the arguments and the verb. No direct evidence was obtained for learners' construction of structurally present gaps, but the findings are not inconsistent with the assumption that phrase structure representations, built during a later processing stage than in NSs, might involve such empty categories. In general, the results of the present study do not provide support for the claims of the SSH that late L2 learners differ from NSs in that they underuse syntactic cues during sentence processing. The finding that a subgroup of learners displayed the same ELAN response to the filled gap as the NSs in Hestvik et al.'s (2007) study also challenges the assumption of the SSH that qualitative differences between NSs' and late L2 learners' processing mechanisms persist even when very high proficiency level has been achieved; the results of this study imply that late learners of English can attain native-like sentence processing skills.

The conclusion of this study is that sentence processing might be slower in L2 learners than in NS, which entails that syntactic phrase structure building operations that have been argued to constitute an initial, autonomous stage in native sentence comprehension, might take place later, parallel with or following other semantic and syntactic operations. The implication of this suggestion is that various syntactic and semantic information types can potentially influence the operation of these processes, indicating that learners might not construct a phrase structure representation based solely on the legal combination of word

categories. This, however, does not mean that L2 learners differ from NSs in their ability to exploit different information sources as suggested by the SSH. Nevertheless, the way these information types interact might be different due to essentially quantitative differences in processing speed. The fact that the present study did not involve a comparison between L2 learners' and NSs' ERP responses obtained under identical experimental conditions imposes limitations on the conclusions that can be drawn. Still, the results are indicative of tendencies that are worthy of further research. Studies making a direct comparison between L2 learners and NSs should investigate the temporal pattern and the nature of interaction of various sentence processing mechanisms in L2 and native language comprehension. In addition, future studies should clarify the role that different factors such as speed of lexical access, phonetic decoding, grammatical knowledge and working memory span play in accounting for the slower sentence processing speed of L2 learners; a limitation of the present study is that the participants' language proficiency was not thoroughly tested, and it is not possible to identify what distinguishes the learner subgroups defined on the basis of processing speed. Finally, it remains to be seen whether the assumption that the observed characteristics of L2 sentence comprehension are related to slower processing is tenable with different proficiency levels and with a variety of sentence structures.

References

Aoshima, S., Phillips, C., & Weinberg, A. S. (2004). Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language, 51,* 23–54.

Bialystok, E., & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 161–181). Mahwah, NJ: Erlbaum.

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning 56,* 9–49.

Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L., Genç, S., Philipp, M., Roehm, D., & Schlesewsky, M. (2011). Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain and Language, 117*, 133-152.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27,* 3-42.

Clifton, C., & Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In G. M. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 273–317). Dordrecht: Kluwer Academic.

DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition 22,* 499–533.

Dikker, S., Rabagliati, H., & Pylkkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition, 110,* 293–321.

Felser, C., & Clahsen, H. (2009). Grammatical processing of spoken language in child and adult language learners. *Journal of Psycholinguistic Research, 38,* 305-319.

Felser, C., Clahsen, H., & Münte, T. F. (2003). Storage and integration in the processing of filler-gap dependencies: An ERP study of topicalization and *wh-movement* in German. *Brain and Language, 87,* 345–354.

Felser, C., Cunnings, I., Batterham, C., & Clahsen, H. (2012). The timing of island effects in nonnative sentence processing. *Studies in Second Language Acquisition 34,* 67-98.

Felser, C., Marinis, T., & Clahsen, H. (2003). Children's processing of ambiguous sentences: A study of relative clause attachment. *Language Acquisition, 11,* 127–163.

Felser, C., & Roberts, L. (2007). Processing wh-dependencies in a second language: a cross modal priming study. *Second Language Research, 23, 9–36.*

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2001). Syntactic working memory and the establishment of filler-gap dependencies: Insights from ERPs and fMRI. *Journal* of *Psycholinguistic Research, 30,* 321-338.

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-

questions. *Journal of Memory and Language, 47,* 250–272.

Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America 106,* 2973–2987.

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology, 50,* 119–148.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences , 6,* 78-84.

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews, 91,* 1357-1392.

Friederici, A. D., & Frisch, S. (2000). Verb argument structure processing: the role of verb-specific and argument-specific information. *Journal of Memory and Language, 43,* 476–507.

Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). The temporal structure of syntactic parsing: early and late ERP effects elicited by syntactic anomalies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1219–1248.

Friederici, A. D., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research*, *31,* 45-63.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological, and syntactic violations. *Cognitive Brain Research, 1,* 183–192.

Frisch, S., Hahne, A., & Friederici, A. D. (2004). Word category and verb-argument structure information in the dynamics of parsing. *Cognition, 91,* 191-219.

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

Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., & Hickock, G. (1996). Recency preferences in the human sentence processing mechanism. *Cognition, 59,* 23–59.

Hagoort P., & Brown, C. (2000). ERP effects of listening to speech compared to reading: The P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia, 38,* 1531–1549.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8,* 439–483.

Hahne, A. (2001). What's different in second-language processing? Evidence from event related brain potentials. *Journal of Psycholinguistic Research, 30,* 251-266.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11,* 193–204.

Hahne, A., & Friederici, A. D. (2001). Processing a second language: late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition, 4,* 123-141.

Herrmann, B., Maess, B., Hasting, A. S., & Friederici, A. D. (2009). Localization of the syntactic mismatch negativity in the temporal cortex: an MEG study, *Neuroimage, 48*, 590–600.

Hestvik, A., Bradley, C., Bradley, E., Kaufmann, M., Prescott, T., & Sparacino, L. (2009). *An event-related potentials measure of the effect of low verbal memory span on gap-filling.* Paper presented at the 22nd Annual CUNY Conference on Human Sentence Processing, Davis, CA.

Hestvik, A., Maxfield, N., Schwartz, R., & Shafer, V. (2007). Brain responses to filled gaps. *Brain and Language, 100,* 301-316.

Hinojosa, J. A., Martin-Loeches, M., Casado, P., Munoz, F., & Rubia, F. (2003). Similarities and differences between phrase structure and morphosyntactic violations in Spanish: an event-related potentials study. *Language and Cognitive Processes, 18,* 113–142.

Hoeks, J. C. J., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research, 19,* 59–73.

Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research, 22,* 369–397.

Isel, F**.** (2007). Syntactic and referential processes in second-language learners: Event-related brain potential evidence. *Neuroreport, 18,* 1885-1889.

Isel, F., Hahne, A., Maess, B., & Friederici, A. D. (2007). Neurodynamics of sentence interpretation: ERP evidence from French. *Biological Psychology, 74(3),* 337–346.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21,* 60-99.

Juffs, A. (2005). The influence of first language on the processing of *wh-movement* in English as a second language. *Second Language Research, 21,* 121–151.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes, 15,* 159–201.

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52,* 205–225.

Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing

of unbounded dependencies. *Journal of Cognitive Neuroscience, 5,* 196–214.

Kolk, H. H. J., Chwilla, D. J., Van Herten, M., & Oor, P. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language, 85,* 1–36.

Kuperberg, G. R., Caplan, D., Sitnikova, T., Eddy, M., & Holcomb, P. (2006). Neural correlates of processing syntactic, semantic and thematic relationships in sentences. *Language and Cognitive Processes, 21,* 489–530.

Kuperberg, G. R., Sitnikova, T., Caplan, P., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptuarelationships within simple sentences. *Cognitive Brain Research, 17,* 117–129.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4,* 463-470.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology, 62,* 621-647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science, 207,* 203–205.

Kutas, M., & van Petten, C. K. (2006). Psycholinguistics electrified II (1994-2005). In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 61–91). Amsterdam: Elsevier.

Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language, 98,* 74–88.

Love, T., & Swinney, D. (1997). *Real-time processing of object-relative constructions by pre-school children.* Paper presented at the 10th annual CUNY conference on human sentence processing. Los Angeles, CA.

Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition, 27,* 53-78.

Martin-Loeches M., Munoz F., Casado P., Melcon A., & Fernandez-Frias, C. (2005). Are the anterior negativities to grammatical violations indexing working memory? *Psychophysiology, 42,* 508–519.

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language, 55,* 381–401

Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex, 47,* 908-930.

Moreno, E. M., & Kutas, M. (2005). Processing semantic anomalies in two languages: an electrophysiological exploration in both languages of Spanish–English bilinguals. *Cognitive Brain Research, 22,* 205–220

Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2-attainment. *Language Teaching, 44,* 1-35.

Neville, H., Nicol, J., Barss, A., Forster, K. I., & Garrett, M. I. (1991). Syntactically-based sentence processing classes: evidence from event- related brain potentials. *Journal of Cognitive Neuroscience, 3,* 151–165.

Nicol, J. L., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research, 18,* 5–19.

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: Effects of proficiency. *Journal of Cognitive Neuroscience, 17,* 1212–1228.

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20,* 786-803.

Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition, 24,* 501–528.

Phillips, C., Kazanina, N., & Abada, S.H. (2005). ERP effects of the processing of syntactic long- distance dependencies. *Cognitive Brain Research, 22,* 407–428.

Phillips, C., & Wagers, M. (2007). Relating Structure and Time in Linguistics and Psycholinguistics. In G. Gaskell (ed.), *Oxford Handbook of Psycholinguistics* (pp. 739-757). Oxford: Oxford University Press.

Pickering, M. J. (2000). No evidence for traces in sentence comprehension (Commentary). *Behavioral and Brain Sciences, 23,* 47–48.

Pickering, M. J., & Barry, G. D. (1991). Sentence processing without empty categories. *Language and Cognitive Processes, 6,* 229–259.

Roberts, L., & Felser, C. (2011). Plausibility and recovery from garden paths in second language sentence processing. *Applied Psycholinguistics, 32,* 299-331.

Roberts, L., Marinis, T., Felser, C., & Clahsen, H. (2007). Antecedent priming at gap positions in children's sentence processing. *Journal of Psycholinguistic Research, 36(3),* 175-188.

Rossi, S., Gugler, M. F., Friederici, A. D, & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience, 18,* 2030–2048.

Rossi, S., Gugler, F. M., Hahne, A., & Friederici, A. D. (2005). When word category information encounters morphosyntax: An ERP study. *Neuroscience Letters, 384,* 228–233.

Segalowitz, N. (2003). Automaticity and second language acquisition. In C. Doughty & M. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 382-408). Oxford: Blackwell Publishers.

Service, E., Simola, M., Metsanheimo, O., & Maury, S. (2002). Bilingual working memory span is affected by language skill. *European Journal of Cognitive Psychology, 14,* 383–408.

Singleton, D. (2005). The Critical Period Hypothesis: A coat of many colours. *International Review of Applied Linguistics in Language Teaching 43,* 269–285.

Stowe, L. A. (1986). Evidence for on-line gap-location. *Language and Cognitive Processes, 1,* 227–45.

Stroud, C., & Phillips, C. (2012). Examining the evidence for an independent semantic analyzer: An ERP study in Spanish. *Brain and Language, 120*, 107-126.

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

Ullman, M. T. (2006). The declarative/procedural model and the shallow structure hypothesis. *Applied Psycholinguistics, 27,* 97–105.

van Hell, J. G., & Tokowicz, N. (2010). Event-related brain potentials and second language learning: syntactic processing in late L2 learners at different L2 proficiency levels. *Second Language Research, 26,* 43-74.

van Herten, M., Kolk, H. H. J., & Chwilla, D. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research, 22,* 241–255.

Vos, S. H., Gunter, T. C., Kolk, H. H. J., & Mulder, G. (2001). Working memory constraints on syntactic processing: An electrophysiological investigation. *Psychophysiology, 38,* 41–63.

Wagers, M., & Phillips, C. (2012). *Going the distance: memory and decision-making in active dependency construction.* Manuscript submitted for publication. Retrieved from http://people.ucsc.edu/~mwagers/work.shtml

Weber-Fox, C.M., & Neville, H.J. (1996). Maturational constrains on functional specializations for language processing: ERPs and behavioural evidence in bilingual speakers. *Journal of Cognitive Neuroscience 8,* 231–256.

Williams, J., Möbius, P., & Kim ,C. (2001). Native and non-native processing of English *wh*-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics, 22,* 509–540.

Appendix

List of experimental sentences

1) Grammatical gap condition
2) Filled gap condition
3) Grammatical object condition
4) Semantic anomaly condition
5) Correct subject-relative condition
6) Semantically anomalous subject-relative condition

1	1	The writer that the assistant greeted at the meeting sat down.
	2	The writer that the assistant greeted the teacher at the meeting sat down.
	3	The writer said that the assistant greeted the teacher at the meeting and then sat down.
	4	The writer said that the assistant greeted the freezer at the meeting and then sat down.
	5	The writer that greeted the assistant at the meeting sat down.
	6	The writer that greeted the banana at the meeting sat down.
2	1	The journalist that the manager tapped on the shoulder walked away.
	2	The journalist that the manager tapped the singer on the shoulder walked away.
	3	The journalist said that the manager tapped the singer on the shoulder and then walked away.
	4	The journalist said that the manager tapped the water on the shoulder and then walked away.
	5	The manager that tapped the journalist on the shoulder walked away.
	6	The manager that tapped the gallery on the shoulder walked away.
3	1	The engineer that the doctor hugged in the corridor opened the window.
	2	The engineer that the doctor hugged the secretary in the corridor opened the window.
	3	The engineer said that the doctor hugged the secretary in the corridor and then opened the window.
	4	The engineer said that the doctor hugged the elevator in the corridor and then opened the window.
	5	The engineer that hugged the doctor in the corridor opened the window.
	6	The engineer that hugged the basement in the corridor opened the window.
4	1	The receptionist that the painter scared by accident answered the phone.
	2	The receptionist that the painter scared the reporter by accident answered the phone.
	3	The receptionist said that the painter scared the reporter by accident and then answered the phone.
	4	The receptionist said that the painter scared the document by accident and then answered the phone.
	5	The receptionist that scared the painter by accident answered the phone.
	6	The receptionist that scared the freezer by accident answered the phone.
5	1	The architect that the accountant met at the restaurant left by car.
	2	The architect that the accountant met the baker at the restaurant left by car.
	3	The architect said that the accountant met the baker at the restaurant and then left by car.
	4	The architect said that the accountant met the apple at the restaurant and then left by car.
	5	The accountant that met the architect at the restaurant left by car.
	6	The accountant that met the tomato at the restaurant left by car.
6	1	The actor that the architect surprised with a gift spilled his drink.

	2	The actor that the architect surprised the translator with a gift spilled his drink.
	3	The actor said that the architect surprised the translator with a gift and then spilled his drink.
	4	The actor said that the architect surprised the windowpane with a gift and then spilled his drink.
	5	The actor that surprised the architect with a gift spilled his drink.
	6	The actor that surprised the cardigan with a gift spilled his drink.
7	1	The coach that the musician invited to the party asked for a pen.
	2	The coach that the musician invited the waitress to the party asked for a pen.
	3	The coach said that the musician invited the waitress to the party and then asked for a pen.
	4	The coach said that the musician invited the napkin to the party and then asked for a pen.
	5	The musician that invited the coach to the party asked for a pen.
	6	The musician that invited the shelf to the party asked for a pen.
8	1	The politician that the pilot thanked for the advice made a phone call.
	2	The politician that the pilot thanked the librarian for the advice made a phone call.
	3	The politician said that the pilot thanked the librarian for the advice and then made a phone call.
	4	The politician said that the pilot thanked the cemetery for the advice and then made a phone call.
	5	The politician that thanked the pilot for the advice made a phone call.
	6	The politician that thanked the napkin for the advice made a phone call.
9	1	The designer that the judge visited in the morning bought a newspaper.
	2	The designer that the judge visited the carpenter in the morning bought a newspaper.
	3	The designer said that the judge visited the carpenter in the morning and then bought a newspaper.
	4	The designer said that the judge visited the banana in the morning and then bought a newspaper.
	5	The designer that visited the judge in the morning bought a newspaper.
	6	The designer that visited the fork in the morning bought a newspaper.
10	1	The editor that the programmer helped with the shopping booked a plane ticket.
	2	The editor that the programmer helped the optician with the shopping booked a plane ticket.
	3	The editor said that the programmer helped the optician with the shopping and then booked a plane ticket.
	4	The editor said that the programmer helped the uniform with the shopping and then booked a plane ticket.
	5	The programmer that helped the editor with the shopping booked a plane ticket.
	6	The programmer that helped the pyramid with the shopping booked a plane ticket.
11	1	The economist that the chemist stopped for a moment called a taxi.
	2	The economist that the chemist stopped the sailor for a moment called a taxi.
	3	The economist said that the chemist stopped the sailor for a moment and then called a taxi.
	4	The economist said that the chemist stopped the apron for a moment and then called a taxi.
	5	The economist that stopped the chemist for a moment called a taxi.
	6	The economist that stopped the ceiling for a moment called a taxi.
12	1	The surgeon that the secretary called in the afternoon went for a run.
	2	The surgeon that the secretary called the bookseller in the afternoon went for a run.
	3	The surgeon said that the secretary called the bookseller in the afternoon and then went for a run.
	4	The surgeon said that the secretary called the strawberry in the afternoon and then went for a run.
	5	The secretary that called the surgeon in the afternoon went for a run.
	6	The secretary that called the water in the afternoon went for a run.

13	1	The referee that the director stopped in the street entered the building.
	2	The referee that the director stopped the biologist in the street entered the building.
	3	The referee said that the director stopped the biologist in the street and then entered the building.
	4	The referee said that the director stopped the macaroni in the street and then entered the building.
	5	The referee that stopped the director in the street entered the building.
	6	The referee that stopped the windowpane in the street entered the building.
14	1	The teacher that the author contacted for information opened his notebook.
	2	The teacher that the author contacted the pharmacist for information opened his notebook.
	3	The teacher said that the author contacted the pharmacist for information and then opened his notebook.
	4	The teacher said that the author contacted the trousers for information and then opened his notebook.
	5	The author that contacted the teacher for information opened his notebook.
	6	The author that contacted the bracelet for information opened his notebook.
15	1	The soldier that the journalist received with politeness took a seat.
	2	The soldier that the journalist received the scientist with politeness took a seat.
	3	The soldier said that the journalist received the scientist with politeness and then took a seat.
	4	The soldier said that the journalist received the living room with politeness and then took a seat.
	5	The soldier that received the journalist with politeness took a seat.
	6	The soldier that received the balcony with politeness took a seat.
16	1	The dentist that the photographer watched from a distance went into the store.
	2	The dentist that the photographer watched the clerk from a distance went into the store.
	3	The dentist said that the photographer watched the clerk from a distance and then went into the store.
	4	The dentist said that the photographer watched the luck from a distance and then went into the store.
	5	The dentist that watched the photographer from a distance went into the store.
	6	The dentist that watched the macaroni from a distance went into the store.
17	1	The singer that the politician kissed on the cheek took off his coat.
	2	The singer that the politician kissed the interpreter on the cheek took off his coat.
	3	The singer said that the politician kissed the interpreter on the cheek and then took off his coat.
	4	The singer said that the politician kissed the motivation on the cheek and then took off his coat.
	5	The politician that kissed the singer on the cheek took off his coat.
	6	The politician that kissed the table on the cheek took off his coat.
18	1	The bookseller that the waitress answered with a nod looked for an empty table.
	2	The bookseller that the waitress answered the salesman with a nod looked for an empty table.
	3	The bookseller said that the waitress answered the salesman with a nod and then looked for an empty table.
	4	The bookseller said that the waitress answered the bracelet with a nod and then looked for an empty table.
	5	The waitress that answered the bookseller with a nod looked for an empty table.
	6	The waitress that answered the tower with a nod looked for an empty table.
19	1	The mechanic that the postman followed through the hall dropped his keys.
	2	The mechanic that the postman followed the cashier through the hall dropped his keys.

3 The mechanic said that the postman followed the cashier through the hall and then dropped his keys.
4 The mechanic said that the postman followed the balcony through the hall and then dropped his keys.
5 The postman that followed the mechanic through the hall dropped his keys.
6 The postman that followed the document through the hall dropped his keys.

20 1 The technician that the consultant joined for lunch ordered a cup of tea.
2 The technician that the consultant joined the driver for lunch ordered a cup of tea.
3 The technician said that the consultant joined the driver for lunch and then ordered a cup of tea.
4 The technician said that the consultant joined the pavement for lunch and then ordered a cup of tea.
5 The technician that joined the consultant for lunch ordered a cup of tea.
6 The technician that joined the metronome for lunch ordered a cup of tea.

21 1 The painter that the actor visited after lunch left for a meeting.
2 The painter that the actor visited the musician after lunch left for a meeting.
3 The painter said that the actor visited the musician after lunch and then left for a meeting.
4 The painter said that the actor visited the parachute after lunch and then left for a meeting.
5 The actor that visited the painter after lunch left for a meeting.
6 The actor that visited the pavement after lunch left for a meeting.

22 1 The scientist that the postman attacked from behind sat down on a bench.
2 The scientist that the postman attacked the dentist from behind sat down on a bench.
3 The scientist said that the postman attacked the dentist from behind and then sat down on a bench.
4 The scientist said that the postman attacked the ceiling from behind and then sat down on a bench.
5 The postman that attacked the scientist from behind sat down on a bench.
6 The postman that attacked the sanity from behind sat down on a bench.

23 1 The chef that the reporter contacted after the meeting made a note in his diary.
2 The chef that the reporter contacted the editor after the meeting made a note in his diary.
3 The chef said that the reporter contacted the editor after the meeting and then made a note in his diary.
4 The chef said that the reporter contacted the metronome after the meeting and then made a note in his diary.
5 The reporter that contacted the chef after the meeting made a note in his diary.
6 The reporter that contacted the floor after the meeting made a note in his diary.

24 1 The gardener that the assistant met by the entrance looked at the advertisements.
2 The gardener that the assistant met the electrician by the entrance looked at the advertisements.
3 The gardener said that the assistant met the electrician by the entrance and then looked at the advertisements.
4 The gardener said that the assistant met the thermometer by the entrance and then looked at the advertisements.
5 The assistant that met the gardener by the entrance looked at the advertisements.
6 The assistant that met the trousers by the entrance looked at the advertisements.

25 1 The author that the accountant received in his office signed the contract.
2 The author that the accountant received the manager in his office signed the contract.
3 The author said that the accountant received the manager in his office and then signed the contract.

4 The author said that the accountant received the tower in his office and then signed the contract.
5 The author that received the accountant in his office signed the contract.
6 The author that received the satellite in his office signed the contract.

26 1 The driver that the librarian helped with the application filled in the form.
2 The driver that the librarian helped the cook with the application filled in the form.
3 The driver said that the librarian helped the cook with the application and then filled in the form.
4 The driver said that the librarian helped the shelf with the application and then filled in the form.
5 The driver that helped the librarian with the application filled in the form.
6 The driver that helped the thermometer with the application filled in the form.

27 1 The salesman that the electrician tapped on the back wiped his glasses.
2 The salesman that the electrician tapped the conductor on the back wiped his glasses.
3 The salesman said that the electrician tapped the conductor on the back and then wiped his glasses.
4 The salesman said that the electrician tapped the gallery on the back and then wiped his glasses.
5 The salesman that tapped the electrician on the back wiped his glasses.
6 The salesman that tapped the cemetery on the back wiped his glasses.

28 1 The biologist that the clerk followed into the room closed the curtains.
2 The biologist that the clerk followed the surgeon into the room closed the curtains.
3 The biologist said that the clerk followed the surgeon into the room and then closed the curtains.
4 The biologist said that the clerk followed the socket into the room and then closed the curtains.
5 The biologist that followed the clerk into the room closed the curtains.
6 The biologist that followed the luck into the room closed the curtains.

29 1 The cashier that the baker attacked in his sleep turned on the lights.
2 The cashier that the baker attacked the gardener in his sleep turned on the lights.
3 The cashier said that the baker attacked the gardener in his sleep and then turned on the lights.
4 The cashier said that the baker attacked the bulldozer in his sleep and then turned on the lights.
5 The baker that attacked the cashier in his sleep turned on the lights.
6 The baker that attacked the strawberry in his sleep turned on the lights.

30 1 The sailor that the pharmacist answered with a smile lifted the suitcase.
2 The sailor that the pharmacist answered the consultant with a smile lifted the suitcase.
3 The sailor said that the pharmacist answered the consultant with a smile and then lifted the suitcase.
4 The sailor said that the pharmacist answered the tomato with a smile and then lifted the suitcase.
5 The sailor that answered the pharmacist with a smile lifted the suitcase.
6 The sailor that answered the living room with a smile lifted the suitcase.

31 1 The fireman that the referee scared with a mask drank a glass of water.
2 The fireman that the referee scared the writer with a mask drank a glass of water.
3 The fireman said that the referee scared the writer with a mask and then drank a glass of water.
4 The fireman said that the referee scared the table with a mask and then drank a glass of water.
5 The referee that scared the fireman with a mask drank a glass of water.
6 The referee that scared the uniform with a mask drank a glass of water.

32 1 The cook that the photographer greeted with a handshake looked out of the window.
2 The cook that the photographer greeted the judge with a handshake looked out of the window.
3 The cook said that the photographer greeted the judge with a handshake and then looked out of the window.

	4	The cook said that the photographer greeted the floor with a handshake and then looked out of the window.
	5	The photographer that greeted the cook with a handshake looked out of the window.
	6	The photographer that greeted the roof with a handshake looked out of the window.
33	1	The mechanic that the conductor thanked for the ride glanced at his watch.
	2	The mechanic that the conductor thanked the technician for the ride glanced at his watch.
	3	The mechanic said that the conductor thanked the technician for the ride and then glanced at his watch.
	4	The mechanic said that the conductor thanked the pyramid for the ride and then glanced at his watch.
	5	The conductor that thanked the mechanic for the ride glanced at his watch.
	6	The conductor that thanked the parachute for the ride glanced at his watch.
34	1	The engineer that the pilot joined for a picture waved at the girl.
	2	The engineer that the pilot joined the programmer for a picture waved at the girl.
	3	The engineer said that the pilot joined the programmer for a picture and then waved at the girl.
	4	The engineer said that the pilot joined the cardigan for a picture and then waved at the girl.
	5	The engineer that joined the pilot for a picture waved at the girl.
	6	The engineer that joined the socket for a picture waved at the girl.
35	1	The doctor that the economist kissed to everyone's surprise left in a hurry.
	2	The doctor that the economist kissed the soldier to everyone's surprise left in a hurry.
	3	The doctor said that the economist kissed the soldier to everyone's surprise and then left in a hurry.
	4	The doctor said that the economist kissed the basement to everyone's surprise and then left in a hurry.
	5	The doctor that kissed the economist to everyone's surprise left in a hurry.
	6	The doctor that kissed the elevator to everyone's surprise left in a hurry.
36	1	The interpreter that the lawyer hugged with one arm walked up to the podium.
	2	The interpreter that the lawyer hugged the coach with one arm walked up to the podium.
	3	The interpreter said that the lawyer hugged the coach with one arm and then walked up to the podium.
	4	The interpreter said that the lawyer hugged the roof with one arm and then walked up to the podium.
	5	The interpreter that hugged the lawyer with one arm walked up to the podium.
	6	The interpreter that hugged the apple with one arm walked up to the podium.
37	1	The butcher that the chemist invited for the interview submitted the report.
	2	The butcher that the chemist invited the director for the interview submitted the report.
	3	The butcher said that the chemist invited the director for the interview and then submitted the report.
	4	The butcher said that the chemist invited the satellite for the interview and then submitted the report.
	5	The chemist that invited the butcher for the interview submitted the report.
	6	The chemist that invited the apron for the interview submitted the report.
38	1	The translator that the optician called at noon filed a complaint.
	2	The translator that the optician called the chef at noon filed a complaint.
	3	The translator said that the optician called the chef at noon and then filed a complaint.
	4	The translator said that the optician called the fork at noon and then filed a complaint.
	5	The optician that called the translator at noon filed a complaint.

	6	The optician that called the detergent at noon filed a complaint.
39	1	The carpenter that the butcher surprised at work took a long walk.
	2	The carpenter that the butcher surprised the designer at work took a long walk.
	3	The carpenter said that the butcher surprised the designer at work and then took a long walk.
	4	The carpenter said that the butcher surprised the detergent at work and then took a long walk.
	5	The butcher that surprised the carpenter at work took a long walk.
	6	The butcher that surprised the bulldozer at work took a long walk.
40	1	The lawyer that the receptionist watched for an hour lighted a cigarette.
	2	The lawyer that the receptionist watched the fireman for an hour lighted a cigarette.
	3	The lawyer said that the receptionist watched the fireman for an hour and then lighted a cigarette.
	4	The lawyer said that the receptionist watched the sanity for an hour and then lighted a cigarette.
	5	The lawyer that watched the receptionist for an hour lighted a cigarette.
	6	The lawyer that watched the motivation for an hour lighted a cigarette.