Syntactic complexity and ambiguity resolution in a free word order language: behavioral and electrophysiological evidences from Basque

Kepa Erdocia[1, 2, 3]

Itziar Laka[4]

Anna Mestres-Missé[5]

Antoni Rodriguez-Fornells[5,6]

[1] INSERM, Unité 562, F-91191 Gif-sur-Yvette, France

[2] CEA, DSV, I2BM, NeuroSpin, F-91191 Gif-sur-Yvette, France

[3] Université Paris-Sud, F-91405 Orsay, France

[4] University of the Basque Country UPV/EHU

[5] Faculty of Psychology, University of Barcelona, 08035, Barcelona, Spain

[6] Institució Catalana de Recerca i Estudis Avançats (ICREA)

Correspondence:
Kepa Erdocia
CEA/NEUROSPIN
Bât 145
91191 GIF/YVETTE
FRANCE
Phone: +33 (0)169867832
Fax: +33 (0)169087973
Email: kepa.erdocia@cea.fr

Key Words: Linguistic complexity, word order processing, Syntactic processing, ERPs, canonicity, Basque, ergativity.

Abstract

In natural languages some syntactic structures are simpler than others. Syntactically complex structures require further computation that is not required by syntactically simple structures. In particular, canonical, basic word order represents the simplest sentence structure. Natural languages have different canonical word orders, and they vary in the degree of word order freedom they allow. In the case of free word order, whether canonical word order plays any role in processing is still unclear. In this paper, we present behavioral and electrophysiological evidence that simpler, canonical word order preference is found even in a free word order language.

Canonical and derived structures were compared in two self-paced reading and one ERPs experiment. Non-canonical sentences required further syntactic computation in Basque, they showed longer reading times and a modulation of anterior negativities and P600 components providing evidence that even in free word order, case-marking grammars, underlying canonical word order can play a relevant role in sentence processing. These findings could signal universal processing mechanisms because similar processing patterns are found in typologically very distant grammars.

We also provide evidence from syntactically fully ambiguous sequences. Our results on ambiguity resolution showed that fully ambiguous sequences were processed as canonical sentences. Moreover, when fully ambiguous sequences were forced to complex interpretation by means of the world knowledge of the participants, a frontal negativity distinguished simple and complex ambiguous sequences. Thus the preference of simple structures is presumably a universal design property for language processing, despite differences on parametric variation of a given grammar.

Introduction

In recent years, a rapidly growing body of experimental studies using neuroimaging techniques has explored syntactic processing of natural language. As a result, findings from linguistics and neuroscience are progressively reaching increasing levels of convergence and reciprocal relevance. However, the vast majority of language neuroimaging studies focus on rather similar languages, such as English, Italian, French, German, Spanish or Dutch. These languages belong to the Indo-European family, and share many central design properties (Baker, 2001). In linguistics, a significant expansion of the language pool investigated was crucial to uncover the interplay between universal and variable aspects of the language faculty (Greenberg, 1963; Chomsky, 1981). It is to be expected, therefore, that the exploration of language in the brain will also benefit from cross-linguistic research, so that we can differentiate language-particular processing strategies from universal language processing mechanisms. In order to discriminate between the two, it is necessary to conduct studies and gather evidence from a wide array of languages pertaining to various typological groups (Bates et al., 2001; Bornkessel and Schlesewsky, 2006).

The present study attempts to broaden the empirical basis of experimental studies, and it does it by presenting and discussing a series of behavioral and electrophysiological experiments in Basque, a free word order, case-marking and ergative language (Laka, 1996; Hualde and Ortiz de Urbina, 2003, De Rijk, 2007). From the large array of potentially relevant features of Basque which could be relevant for the emerging field of neurocognition of language, this paper focuses on two aspects: the relevance of underlying, canonical word order in sentence processing and its role in sentence-ambiguity resolution.

1. Word order processing

Languages vary with respect to word order freedom; some have a very fixed word order, like English, others allow a greater variation of word order, like Spanish, and still others allow for almost all possible combinations of phrases in a sentence, like Basque (Baker 2001). Despite this variation, it has been argued in linguistics that grammars have an underlying, canonical word order (Greenberg 1963, Chomsky 1981), which surfaces in a declarative sentence that initiates discourse, that is, a sentence where no constituent is focalized and where the entire event constitutes new information (Lambrecht 1994). Canonical word order thus reflects the simplest phrase sequence generated by the grammar. Most human languages, independently of the degree of word order freedom, group into two main types regarding canonical word order (Greenberg, 1963): *Subject-Verb-Object* (SVO) languages, such as English or Spanish, and *Subject-Object-Verb* (SOV) languages, such as Japanese or Basque (these two main types are also known as *head-first* and *head-last* languages within the Principles and Parameters model, Chomsky 1981, Baker 2001).

A wide array of experiments have explored the asymmetries between the subjects and objects in relative clauses (e.g. King & Kutas, 1995; Müller et al., 1997; Gibson 1998, Traxler et al. 2002), in interrogative sentences (e.g. Fiebach et al., 2002; Felser et al., 2003; Bornkessel et al., 2004, Ben-Shachar et al., 2004) and in declarative sentences (e.g. Rosler et al., 1998; Matzke et al., 2002; Balhmann et al., 2007; Schelewsky et al., 2003; Hagiwara et al., 2007). Most of these researches found that *object-before-subject* structures require higher processing effort than *subject-before-object* structures. For instance, subject relative clauses (*The reporter [who attacked the senator] admitted the error*) and subject interrogative clauses (*Thomas asks himself [who called the doctor]*) are easier to process than object relatives (*The reporter [who the senator attacked] admitted the error*) and object interrogatives (*Thomas asks himself [who the doctor called]*). These studies have proposed that the larger processing effort observed might be due either to the syntactic complexity of the transformations or to the load in working memory (see Fiebach et al. 2005).

However, Traxler et al. (2002) showed that the processing cost of the object relatives with respect to the subject relatives could be reduced manipulating the semantic properties of the subjects and objects. For example, this reduction has been accomplished manipulating the animacy of the subjects and the objects of subject relatives and object relatives (Mak et al., 2006), manipulating the argument structure of the verbs (Bornkessel et al., 2005; Schlewsky and Bornkessel, 2006), or comparing pronominal subjects and objects (Schlesewsky et al., 2003). These results could not be explained by the transformation-based framework or by the working memory framework. Bornkessel et al. (2005) argued that the higher processing effort of the sentences could be related to the syntactic-semantic interface. These authors suggested that the linearization of the constituents' thematic role in a given sentence could vary the processing effort of the sentences independently or jointly of the syntactic structure of the sentence (Bornkessel et al., 2005).

Regarding to the signs of the processing effort, behaviorally, easier processing correlates with shorter reading times and/or shorter error rates (see Sekerina, 2003, for an overview). Electrophysiological studies, on the other hand, showed a variety of results. For instance, Object interrogative clauses showed a left anterior negativity (LAN) and P600 pattern compared with Subject interrogatives (Fiebach et al., 2002; Felser et al., 2003). This pattern has been explained as the storage and the integration of displaced elements respectively (Fiebach et al., 2002; Felser et al., 2003). On the other hand, Bornkessel et al. (2004) showed a sustained negativity in the Object interrogatives and a N400 in the object-second position of Subject interrogatives. This N400 effect could be explained appealing to the Minimality-driven processing strategy (Bornkessel & Schlesewsky, 2006) which predicts that the subject-initial sentences are considered intransitive in German, and processing the unpredictable object then generates a N400 related to the semantic integration. Thus, in some ERP studies the subject/object asymmetries in non-declarative sentences are considered to be syntactic processes because the LAN and the P600 components are related to syntactic processing (Fiebach et al., 2002; Felser et al., 2003). In other studies, the N400 component related the subject/object asymmetries with the processing of argument structure (Bornkessel et al., 2004).

In declarative sentences, behavioral studies have shown that sentence processing is sensitive to word order. In most languages canonical word order is processed faster and with greater ease (see Sekerina, 2003, for an overview). However, in a scrambling language like Japanese, the results of behavioral studies do not converge. While some self-paced reading studies reported no differences between SOV and OSV word orders in Japanese (Yamashita, 1997; Tamaoka et al., 2003), other studies showed that OSV word orders required higher processing demands (Miyamoto and Takahashi, 2002; Mazuka et al., 2002).

ERP studies have investigated word order processing in sentences in which no interrogative or relative marker is heading the clause. Most of these studies were carried out in German (Rösler et al., 1998; Matzke et al., 2002; Bornkessel et al. 2002, Schlesewsky et al., 2003). German is not a fixed word order language: it allows SVO word order as in the sentence "*der Mann sah den Jungen*" (the man-SUBJECT saw the boy-OBJECT) but it also allows other word orders like OVS as in, "*den Jungen sah der Mann*" (the boy-OBJECT saw the man-SUBJECT). The ERP studies of Rösler et al. (1998) and Matzke et al. (2002) observed a LAN component at object-first position, which has been attributed to the storage cost of the displaced Object-phrase. The authors argued that the displaced Object-phrase must wait until its canonical position is reached for interpretation. This result agrees with previous studies in which differences in word order using relative or interrogative clauses were explored. In these studies, the LAN component has been attributed to syntactic working memory storage (Kluender and Kutas, 1993; King and Kutas, 1995; Müller et al., 1997; Münte et al., 1998; Fiebach et al., 2002; Felser et al., 2003). In a recent functional magnetic resonance imaging (fMRI) study, and using the same type of sentences used by Matzke et al. (2002), Bahlmann et al. (2007) reported that non-canonical sentences elicited larger activation in the left inferior frontal gyrus. This result confirmed the idea that non-canonical sentences require a larger integration cost and make greater demands on working memory. These results in turn also converge with linguistic accounts of German syntax, where object-initial main sentences involve displacing the object from its canonical position to a higher place in the sentence-structure (Schwartz and Vikner, 1996). On the other hand, comparing sentences initial Accusative objects and Nominative Subjects, Bornkessel et al (2002) showed a central negativity between 300 and 450 milliseconds, but comparing sentences initial Dative objects and Nominative subjects no differences were observed. These results converged with the results of an fMRI study in which verb argument demands required dative objects (Bornkessel et al., 2005), demonstrating that the syntactic complexity can be modulated by means of the semantic properties of the verbs (active vs. object-experiencer) and the semantic properties of nouns (animate vs. inanimate).

In contrast to these previous ERP studies on German, a recent study in Japanese did not replicate the presence of a negative component in object-first position of non-canonical declarative sentences (Hagiwara et al., 2007). In Japanese, canonical word order is SOV as in the sentence "*hishoga bengoshio sagasiteiru*" (the secretary-SUBJECT the lawyer-OBJECT was looking for). However, it also allows other verb-final word orders, even to a greater degree than German, as for example in the sentence "*bengoshio hishoga sagasiteiru*" (the lawyer-OBJECT the secretary-SUBJECT was looking for). Hagiwara et al. (2007) interpreted the lack of LAN in object-first position sentences, as an indication that the non-canonical OSV structure is not complex enough to elicit an ERP component at Object-position (Hagiwara et al., 2007). Nevertheless, when Object-phrase was displaced from an embedded clause, at a greater distance, the complexity increased and the LAN component was observed at object-first position.

Moving along the constituents in the sentence, Matzke et al. (2002) reported a LAN effect in subject position of OVS non-canonical German sentences. Rösler et al. (1998), using German verb-final sentences, also reported a LAN-like component in sentence second position comparing the non-canonical Objects (S-*O*-IO-V) to canonical Indirect Objects (S-*IO*-O-V). Finally, Bornkessel et al. (2002) reported an early positivity (in between 300 and 400 ms) at subject second position in Dative Object initial sentences. In Japanese, the S in an OSV sentence (compared to the O of SOV) showed a Frontal Negativity in the right hemisphere (Hagiwara et al., 2007).

2. Some relevant properties of Basque grammar

Basque (or Euskara, as it is known to its speakers) is an isolate language; it does not belong to any known language family. It is a free word order language; nearly all constituent combinations yield a grammatical sentence, as shown partially in (1):

(1)a. emakume-a-k gizon-a ikusi du gaur

woman-the-erg man-the seen has today

'the woman has seen the man today'

b. gizona ikusi du gaur emakumeak

c. gizona ikusi du emakumeak gaur

d. gaur ikusi du emakumeak gizona

e. gaur ikusi du gizona emakumeak

f. emakumeak ikusi du gizona gaur

g. emakumeak ikusi du gaur gizona

h. gizona emakumeak ikusi du gaur

i. gizona gaur ikusi du emakumeak

j. gaur gizona emakumeak ikusi du

k. ikusi du emakumeak gizona gaur

Despite this freedom in sentence word order, Basque grammar has been argued to be of the SOV type (Greenberg 1963, de Rijk 1969), because it has most correlated properties of this type: relative clauses and genitives are prenominal (2a, b), it has postpositions and suffixes instead of prepositions (1, 2b), determiners follow the noun (1), (2a, b), and inflected auxiliaries follow the verb (1), (2a):

(2)a. [emakume-a-k e_1 ikusi du-en] gizon-a_1

[woman-the-erg e_1 seen has-rel] man-the

'the man that the woman has seen'

b. [emakume-a-ren] lagun-a

woman-the-gen friend-the

'the woman's friend'

Given these properties, linguists consider Basque to be a Head-final language (Ortiz de Urbina 1989, Laka 1994, Baker 2001). Basque is a case-marking language (1), (3), and the verb agrees with the subject, the object and the dative, if contained in the sentence (3a, b).

(3)a. ni-k zu-ri musu-a eman d-i-zu-t

I-erg you-dat kiss-the given it-root-you-me

'I have given you a kiss'

b. zu-k ni ikusi na-u-zu

you-erg me seen me-root-you

'you have seen me'

Basque is three-way pro-drop language (Ortiz de Urbina 1989, Laka 1996): subjects, objects and datives can be phonologically unrealized:

(4)a. eman d-i-zu-t

given it-root-you-me

'I have given it to you'

b. ikusi na-u-zu

seen me-root-you

'you have seen me'

Basque is also an ergative language (Dixon 1994, Levin (1983), Ortiz de Urbina (1989)); this means that subjects of intransitive clauses (S_{INTR}) and objects of transitive clauses (O) are morphologically identical, and bear no overt case ending, while agentive subjects of transitive clauses (S) are morphologically distinct, and carry and ergative case marker (5):

(5)a. gizon-a etorri da

man-the arrived is

'the man arrived'

b. emakume-a-k gizon-a ikusi du

woman-the-erg man-the seen has

'the woman has seen the man'

Given this combination of grammatical features, if the constituent *gizona* "the man" is encountered at the beginning of an utterance, the following possibilities arise: (i) it is the S(ubject) of an SV intransitive sentence like (5a); (ii) it is the O(bject) of an OSV sentence like (1b,c,h,i); (iii) it is the O of a OV transitive sentence where S has been pro-dropped as in (4).

Basque Noun-Phrase morphology presents an interesting homomorphism that has been exploited in this study to explore structural ambiguity resolution. As shown in (6a, b), the form of the plural determiner for either S_{INTR} or O is *–ak*, and as shown in (6b) the combination of the singular determiner *–a* plus the ergative marker *–k* yields a sequence *–ak* that is homophonous with the plural determiner. Consequently, given the free word order property, the same sequence of sounds could be interpreted as a canonical SOV sentence (6b) or as a non-canonical OSV sentence (6c).

(6) a. gizon-ak etorri dira

men-the$_{pl}$ arrived are

"the men arrived"

b. emakume-a-k gizon-ak ikusi ditu

woman-the-erg man-the$_{pl}$ seen has$_{pl}$

"the woman has seen the men"

c. emakume-ak gizon-a-k ikusi ditu

woman-the$_{pl}$ man-the-erg seen have$_{pl}$

"the women, the man has seen them"

In the absence of any other means to disambiguate the sentence, both (6b) and (6c) are possible parsings.

3. Sentence-ambiguity resolution

Ambiguity is a pervasive phenomenon in all dimensions of natural language (semantics, phonology, morphology, syntax), and it has been massively explored in language processing at least since Bever (1970). Frazier and Fodor (1978) assumed that when confronted with a syntactically ambiguous structure, the simplest one is favored. If true, this predicts that canonical word order will be the preferred interpretation given an ambiguous input. Matzke et al. (2002) compared temporally ambiguous SVO sentences (e.g., *die Frau hatte den Mann gesehen;* lit.: the woman-AMBIGUOUS has the man-OBJECT seen) to temporally ambiguous OVS sentences (e.g. *die Frau hatte der Mann gesehen* lit.: the woman-AMBIGUOUS has the man-SUBJECT seen). Notice that in both sentences the initial feminine noun phrases are compatible with either a *subject-first* or an *object-first* interpretation due to the morphological ambiguity of the determiner “die”. The ERP results showed a P600 component at the disambiguation point only for OVS sentences. The authors interpreted this effect as a reanalysis of syntactic structure after an initial subject-first interpretation. Matzke et al. (2002) did not find any difference between unambiguous and ambiguous SVO sentences at the beginning of the sentence, suggesting a general subject-first processing strategy for German (Bates et al., 1988, Schlesewsky et al., 2000). However, using similar materials, Frisch et al. (2002) found a P600 component for subject-initial temporally ambiguous sentences, which they interpreted as an indicator of syntactic ambiguity. Similarly, comparing unambiguously marked subject-initial interrogatives with questions starting with an ambiguous marker, Bornkessel et al. (2004) found that high span participants generated a sustained anterior negativity while low span participants generated a sustained parietal positivity.

The present study was designed to investigate word order effects and ambiguity resolution in Basque, using behavioral (self-paced reading) and electrophysiological techniques (ERPs). We carried out three different experiments: The purpose of Experiment 1 was to measure reading time patterns for SOV versus OSV sentences (see Table 1) in order to determine the canonical processing strategy of Basque speakers. In Experiment 2 we aimed at determining whether fully ambiguous sentences (see Table 1) are processed by means of a default subject-first strategy

(Bates et al., 1988). Finally, Experiment 3 aimed to extend and to evaluate the previous word order effects and the world knowledge guided ambiguity resolution effects using ERPs.

Experiment 1

The main objective of this experiment was to find out whether native speakers of Basque process SOV and OSV sentences differently. According to the majority of syntactic analyses, (Hualde and Ortiz de Urbina, 2003, de Rijk 2007) SOV is the canonical word order, and OSV is derived (see examples of Experiment 1 in Table 1). Using a self-paced reading paradigm, reading times to each word and the whole sentence were studied. We expected sentences featuring an OSV word order to require longer reading times than SOV sentences.

METHOD

Participants

Thirty-three native speakers of Basque volunteered to take part in the experiment, all of them students at the University of the Basque Country (UPV/EHU). The data obtained from 10 participants who made more than 37.5% of errors in the comprehension task were discarded. Thus, data from 23 participants (mean age 25 ± 5 years; 13 women and 10 men) were used for statistical analysis. All participants had normal or corrected to normal vision.

Materials

In this first experiment, stimuli consisted of 32 SOV sentences plus a second set of OSV sentences derived from the original 32 SOV samples (Table 1). Two lists were generated in order to prevent participants from reading both versions of the sentence; that is, if one participant read a SOV sentence, then this participant did not read the corresponding OSV sentence. Thus, the material was counterbalanced across participants. The lexical material of the sentences was controlled in length and frequency by means of *E-Hitz* (Perea et al., 2006), the Basque version of the N-Watch (Davis, 2005). The mean word length and frequency for the nouns used as Objects in the present experiment were 6.8 letters and 70.9 (per million). For Subjects, word length was 6.8 and the mean frequency 52.7. No significant differences were encountered when comparing word length and mean frequency between objects and subjects (in both cases, $t < 1$).

Every participant read 16 SOV sentences and 16 OSV ones. In addition, 32 sentences were also included in each list as fillers. Fillers were the same in both sentence-lists. They were matched with the sentences in the experimental conditions in the number of words they contained (e.g., "Manu futbolari bikaina da" *'Manu soccer-player excellent is'*). The experimental session was preceded by 9 warm up practice trials (3 sentences from each condition and 3 fillers) to ensure that participants understood the dynamics of the experiment.

Procedure

Reading times were obtained by presenting the materials to the participants using a *self paced reading* task (Just et al., 1982) in an adapted version of the EXPE v.6 software (Pallier et al., 1997). Determiners and other grammatical elements are attached at the end of the word in Basque (it is an agglutinative, head final language), so that in our materials each word of the sentence constituted a syntactic phrase (i.e. the word "emakumea" is *woman-the*). We presented the sentences word by word (not morpheme by morpheme) by means of a *moving window* technique (Just et al., 1982; Kennedy and Murray, 1984). In this manner, participants read the words as they would normally read them in a written text. In order to move from one word of the sentence to the next, participants had to press the space bar of the computer keyboard recording the time they needed to read each word. The reading times of the participants were measured after the presentation of each word in every sentence. The mean reading time of the whole sentence was computed simply by adding the reading times of each word.

Participants were instructed to read at a normal, comfortable pace and in a way that would enable them to understand the sentences. Each trial began with a series of asterisks marking the length and position of the phrases in the sentences (one asterisk for each letter). Sentences were presented on the computer screen. No punctuation marks were presented throughout the trial. The first pressing of the space bar replaced the first asterisks with the first word of the sentence. As participants pressed the space bar, asterisks were replaced by the next word, while the preceding word was again replaced by asterisks. When participants read the entire sentence, a comprehension question appeared on the screen. The time a participant spent reading each word was recorded as time between key-presses.

The comprehension task allowed us to ensure that participants had understood the sentences they read. The task consisted in answering a yes/no question after each sentence. The answer to half of the questions of each condition was "yes" and the other half was "no". For example, the question corresponding to the test items of Experiment 1 (examples 7a and 7b in Table 1) was: "egia al da emakume batek gizon bat ikusi duela?" *'is it true that a woman has seen a man?'*; in this case the correct answer was "bai" *'yes'*. To answer the question participants had to press the button corresponding to 1 or 2 on the keyboard. Participants did not receive feedback on their answers. We also recorded reaction times to comprehension questions. The comprehension task was performed also after the filler sentences were read, but filler data were not analyzed.

RESULTS

Comparing the reading times of the sentences (see Figure 1A and Table 2), the first experiment revealed that SOV sentences were processed significantly faster than OSV sentences. As it is shown in Table 2, the mean reading time of SOV sentences was significantly faster than for OSV sentences (3755 ms vs. 4151 ms, respectively). Moreover, the comprehension task revealed that answering a question about an OSV sentence was harder than answering a question about a SOV sentence, as evidenced by a significant longer reaction times to provide an answer (see Figure 1B) and a significant larger percentage of errors produced

(Figure 1C and Table 2).[1] Thus, at least under this experimental situation, the processing of OSV sentences was harder than the processing of SOV sentences.

Regarding the word by word reading time analysis (see Figure 1D and Table 2), the results showed significant differences across experimental conditions. When we consider sentence initial words, Subjects (S) required longer reading times than Objects (O) (847 ms vs. 789 ms, respectively); that is, the O of an OSV sentence was read faster than the S of a SOV sentence (first words of examples 7a and 7b in Table 1). Regarding the second word of the sentence, the second position (S) also required a significant longer processing time (1114 ms) than second position O (989 ms). Therefore, Subjects required longer reading times than Objects regardless of their position in the sentence (first or second). The statistical analysis revealed also a significant interaction between sentence type and the first two words of the conditions ($F(1,22) = 12.9$; $P < .002$).

At the Verb (V) position, reading times increased in both conditions, but they were larger in the OSV than in the SOV sentence (see table 2, SOV 1046 ms and OSV 1241 ms). Finally, at the auxiliary verb reading times decreased in both conditions, possibly because grammatical information and argument structure was already processed given the case morphology on the noun phrases. However, the difference between the mean reading time of auxiliary of the SOV condition (873 ms) and the reading time of the auxiliary of the OSV (1007 ms) remained statistically significant.

Discussion

The first experiment revealed that for native Basque speakers, sentences with SOV word order were easier and faster to process than equivalent sentences with OSV word order. This result is convergent with the claim in linguistics that OSV sentences are syntactically more complex than SOV sentences in this language (Hualde and Ortiz de Urbina, 2003; De Rijk, 2007). Thus, the results from our first experiment are consistent with the predominant view in linguistics regarding word order in Basque, which states SOV as canonical, and all others, including OSV, as structurally more complexes (De Rijk, 1969; Ortiz de Urbina, 1989; Arregi, 2002; Elordieta, 2004).

Secondly, word-by-word analysis showed that Subjects required longer reading times than Objects regardless of their relative position in the sentence (first or second). This result diverges sharply from what is usually obtained with the same method (self-paced reading) in other languages such as Dutch (Kaan, 1997) and Russian (Sekerina, 1997) where Subjects are always processed faster than Objects. This divergent result could be due to one language-particular property: Basque is an *ergative* language, whereby O and non-agentive S belong to the same morphological class, whereas transitive/agentive S (all instances of S in our experiment are transitive/agentive subjects) carries an extra morpheme (-*k* in the case of Basque). This extra morpheme on transitive/agentive S could cause the increase in processing time (see examples in 5a and 5b). Alternatively or jointly, and given this property of the language, when confronted with sentence initial O sentences participants could hypothesize that they were reading a non-agentive S (example 5a), and project a simple intransitive structure where S and V would be the only elements of the structure, and which could in turn yield faster processing of the word. Longer reading times on second position S could be attributed to the fact that this element appears always after an initial O in our materials. If the previous explanation for faster reading times on sentence-initial O is correct, then participants would realize upon reading the second position S that their initial parsing hypothesis was wrong, and reanalysis would ensue, increasing reading times on second position S as a consequence. So, after processing the sentence initial word, if participants have hypothesized that this is a simple intransitive structure, processing S in second position should cause a reanalysis of the structure of the sentence from the intransitive one to the transitive structure where S, O and V would be elements of the new structure. It is precisely the class of *ergative* languages that has called into question the universality of the class of *subjects*, which is well-defined for nominative languages (Dixon, 1994).

Experiment 2

Given our previous results, suggesting that the processing cost of OSV sentences is greater than the one observed in SOV sentences, we wanted to further explore the processing of fully ambiguous sentences. In particular, we wanted to determine whether fully ambiguous sentences were processed resorting to a default *subject-first* strategy, a processing strategy that has been thoroughly reported in the literature for many languages (Bates et al., 1988).

In this experiment we presented three types of sentences. The first two conditions involved SOV and OSV sentences as in Experiment 1, but now noun phrases were plural instead of singular (examples 8a and 8b in Table 1). Notice that the OSV sentence (example 8b) is temporally ambiguous due to the morphological ambiguity of the *-ak* ending: it can be either singular Subject or plural Object (see also examples 6). However, after reading the second unambiguous word, the only possible choice for the first word of the sentence is Object. The reason is that this second word, *-ek*, in Basque, is unambiguously a plural Subject. Then, if the first word of the sentence had been analyzed as a subject, reanalysis ensues. The mean word length and frequency for the nouns used as Objects in the present experiment were 6.6 letters and 51.2 (per million). For Subjects, word length was 6.5 and the mean frequency 43.1. No significant differences were encountered when comparing word length and mean frequency between objects and subjects conditions (in both cases, $t < 1$).

In the third condition we presented *fully ambiguous* sentences (example 8c in Table 1, 6a and 6b). In this type of sentences, both noun phrases carry the ambiguous ending *-ak*, and therefore, either of them could be subject (singular) or object (plural). In this condition, nothing in the morphology of the noun phrases or the verb disambiguates the respective roles of both constituents of the sentence. If participants show preference for one interpretation over the other, this can only be attributed to a default processing mechanism, since there is no other source for disambiguation.

METHOD

Participants

In this experiment 35 native Basque speakers took part as volunteers. All participants who made more than 37.5% of errors in the comprehension task were eliminated from statistical analysis. In total, data from 23 participants was analyzed (3 men and 20 women; mean age 20.4 ± 2.5 years).

Materials and Procedure

We generated 48 sentences. In order to counterbalance the material across participants, sentences were distributed in three lists. Every list had 16 sentences for each condition. Furthermore we created 48 filler sentences, the same for every list. All procedures and techniques were the same as in Experiment 1.[2]

RESULTS

The results obtained in experiment 2 clearly showed that fully ambiguous sentences were processed as having a SOV word order (see Table 3 and Figure 2). In order to evaluate if there were significant differences across conditions, we performed a repeated measures ANOVA introducing the three conditions (SOV, OSV, AMB). This analysis showed a main effect of condition ($F(2,44) = 13.9$, $P < 0.001$). In Figure 2a, it can be observed that the global mean reading time for unambiguous SOV sentences (example 8a in Table 1) and fully ambiguous sentences (example 8c in Table 1) was equivalent (4523 ms vs. 4635 ms, respectively). On the other hand, comparing these two conditions (fully ambiguous and unambiguous SOV) to the OSV condition (example 8b in Table 1), we found that the reading time was significantly longer for the latter condition (mean reading time was 5412 ms). Thus, the comparison of the whole sentence's mean reading times showed that native speakers processed fully ambiguous sentences as SOV, suggesting that they employed a default sentence processing strategy that resorts to canonical word order, that is, they assumed that those fully ambiguous sentences had a SOV order.

The results corresponding to the comprehension task (see Table 3 and Figure 2B), showed that the answers to unambiguous SOV sentences generated fewer errors (5.7 %) than the answers to the OSV (8.6 %) and the fully ambiguous sentences (12.8 %).[3] The results of the corresponding repeated measures ANOVA showed a significant main effect of condition ($F(2,44) = 5.8$, $P < 0.006$). Besides, a significant linear trend was observed ($F(1,22) = 13.8$, $P < 0.001$) showing that the percentage of errors increased linearly across the three conditions (SVO, OSV and AMB; see Table 3, for the pairwise comparisons).

In the word-by-word comparison, when we compared SOV and OSV conditions, the ambiguously marked Object required shorter reading times than the unambiguously marked Subject regardless of the sentences initial or second position (Figure 2C). At sentence initial position the ambiguous O of the OSV condition required a significantly shorter reaction time (Table 3) than the unambiguous S of the SOV condition. At sentence second position, the S required again a significant longer reading time than the O (1471 ms vs. 1269 ms). At the Verb position and at the auxiliary position the reading time increased significantly in the OSV condition (see Table 3 and Figure 2C). Finally, as in Experiment 1, results showed an interaction between first both elements of the sentences and the sentence type ($F(1,22) = 17,137$; $P < .001$).

Turning now to the word-by-word comparison of fully ambiguous sentences (Table 3) and the unambiguous SOV condition, the ambiguously marked element was processed significantly faster (940 ms) than the unambiguously marked S (1018 ms) in sentence initial position (Figure 2C). At second position, constituents of both conditions did not differ, meaning that in both cases they were processed like plural objects (see Table 3). The same result was obtained at verb position (see Table 3). But at auxiliary position, we observed that the auxiliary in fully ambiguous sentences required statistically longer reading times than the auxiliary in the SOV condition (1351 ms vs. 1166 ms), even though the former was shorter (*ditu*) than the later (*dituzte*). Finally, the statistical analysis did not show any significant interaction between the SOV and the fully-AMB conditions.

Discussion

The results of this experiment converge with previous studies in which a processing cost has been observed associated to temporally ambiguous sentences like example 8b in Table 1 (Kaan, 1997). The results also showed that the syntactic disambiguation occurred at sentence second position where unambiguously marked Subject forces a reanalysis of the syntactic structure changing the initial Subject-then-Object interpretation to the O-then-S interpretation.

On the other hand, using fully ambiguous sentences (example 8c in Table 1), which had never been used before in previous experimental studies, our experiment showed that fully-ambiguous sentences were systematically processed as canonical SOV sentences. Although we observed differences in initial positions when we compared fully-ambiguous sentences and canonical SOV sentences, the statistical analysis did not show any significant interaction when considering the reading times for both the first and the second constituents. This result could be explained if we bear in mind that, unlike the case of the OSV condition, no syntactic reanalysis is required for the fully-ambiguous condition if it is processed as a SOV sentence. In other words, since speakers were not forced to break down the structure they were building, they could construct the syntactic structure just as they read the sentence without resorting to reanalysis. This result supports the default *subject-first* processing mechanism for fully ambiguous sentences.

Experiment 3

Using ERPs we aimed to obtain electrophysiological evidence of the syntactic complexity observed in previously presented behavioral experiments 1 and 2. On the one hand, we investigated whether the Object of non-canonical OSV sentences, in comparison with the Subject of canonical SOV sentences (respectively, example 9b and 9a in Table 1) generated any electrophysiological response as it has been previously observed in German (Rösler et al., 1998, Matzke et al., 2002) or not, as in Japanese (Hagiwara et al., 2007).

On the other hand, our ERP experiment sought to study processing mechanisms for syntactically fully-ambiguous sentences (examples 9c and 9d, in Table 1). In order to obtain a reanalysis effect in fully ambiguous sentences, the sentences could be disambiguated at Verb position by means of world knowledge. That is, participants were forced to apply their knowledge of the world in order to choose a plausible interpretation for the syntactically ambiguous sentences. As far as we know this is the first time that syntactically fully-ambiguous sentences are analyzed using ERPs.

METHOD

Participants

Twenty-six native speakers of Basque gave informed consent to participate in the experiment. None of the participants had previous neurological history and had normal or corrected-to-normal vision. The participants were all right handed (Edinburgh Handedness inventory: Oldfield 1971). All participants were paid for their participation. Two participants were excluded from all statistical exploration; one due to excessive eye movement artifacts and another one because his native language was Spanish instead of Basque. We thus analyzed the data from 24 subjects (mean age 26 ± 4.7 years; 8 males). A bilingual Basque-Spanish questionnaire was administered to all participants (adapted from Weber-Fox and Neville, 1996). This questionnaire showed that all participants used Basque more often than Spanish.

Material

Using SOV and OSV sentences, and making use of the ambiguous -*ak* morpheme, we created an array of ambiguous sentences. As we have shown in the experiment 2, when confronted with two constituents featuring the ambiguous -*ak* ending, speakers processed the first as the Subject and the second as the Object but not the other way around. In the present experiment, there were two conditions where both constituents again featured the ambiguous -*ak* ending. Even though these sentences were syntactically ambiguous, only one of the two possible interpretations was consistent with our world knowledge. When participants applied the *subject-object* order, sometimes they got a sentences corresponding to a proposition consistent with their world knowledge, as in example 9c in Table 1, but in other instances, as in example 9d in Table 1, they got a sentence corresponding to a proposition that is not consistent with world-knowledge.

For experiment 3 we created 240 sentences leading to 940 sentences overall from the four different variations to fit the four necessary conditions as in examples 9a, 9b, 9c, and 9d in Table 1. In order to visualize the ERP effects better we introduced a Postpositional Phrase (PP) between the two constituents in every sentence. The PP insertion has at least two objectives: (a) on the one hand, we prevent possible ERP effects contamination across constituents; that is, we avoid effects of the first constituents carrying over to the second constituent; and (b) on the other hand, the PP insertion prevents list interpretations where constituents could be considered as coordinated elements. A sample sentence of experiment 3 where the PP is put in italics is provided: Otsoek *Gorbeia mendiko larretan* ardiak jan dituzte” that means ‘*In the meadow of Gorbeia mountain* the wolves have eaten sheep (pl)’.

Concerning to the frequency and the length of the materials in experiment 3 no differences were observed due to the frequency of subjects and objects (Subjects 78.3 ± 163 (SD); Objects 83.9 ± 151; $t < 1$). However, a significant difference was observed between the length of the subjects and objects (t = 6.6, P < 0.05). The mean length of subjects was 7.1 ± 2.2 letters, and the mean length of objects was 5.9 ± 1.7.

Procedure

With the materials we created 4 lists. Every list had 240 different sentences divided in blocks of eight sentences (30 blocks for list). In every block there were two sentences for each condition. The sentences of each block were mixed randomly every experimental session.

Each ERP session lasted about 45-50 minutes. Participants were told that the main purpose of the experiment was to read carefully the sentences presented and to answer correctly the questions related to the sentences. Experimental trials began with a green asterisk in the middle of the blue screen. The words (in yellow) appeared and disappeared automatically in the middle of the screen until the sentence finished (word duration, 250 ms; stimulus onset asynchrony, 500 ms). Once the first sentence finished participants were allowed to blink (during 3000 ms) and the green asterisk appeared again (for 1500 ms) indicating that a new sentence started. Every 8 sentences a sentence fragment was presented in brown letters and the participants' task was to decide whether or not the fragment had been presented in the any of the preceding eight sentences. This question remained on the screen until a response was given. This task was used in order to control the attention of the participants during the experiment.

Electrophysiological Recording

The ERPs were recorded from the scalp using tin electrodes mounted in an electrocap (Electro-Cap International) and located at 29 standard positions (Fp1/2, Fz, F7/8, F3/4, Fc1/2 Fc5/6, Cz, C3/4, T3/4, Cp1/2, Cp5/6, Pz, P3/4, T5/6, Po1/2, O1/2). EEG data was rereferenced off-line to the mean of the activity at the two mastoid processes. Vertical eye movements were monitored with an electrode at the infraorbital ridge of the right eye. Electrode impedances were kept below 5 kOhm.

The electrophysiological signals were filtered with a bandpass of 0.01-50 Hz (half-amplitude cutoffs) and digitized at a rate of 250 Hz. Trials with base-to-peak electro-oculogram (EOG) amplitude of more than 50 µV, amplifier saturation, or a baseline shift exceeding 200 µV/s were automatically rejected off-line. Percentage of artifact rejection was 6.7%.

Data analysis

Stimulus-locked ERPs were averaged for epochs of 1024 ms starting 100 ms prior to the stimulus. First, two omnibus repeated measures ANOVAs were conducted for the initial evaluation of the stimulus-locked ERP activity in two locations: at Parasagittal locations (PS, which included 5 levels for anterior/posterior factor, Fp1/Fp2, F3/F4, C3/C4, P3/P4, O1/O2) and at Temporal locations (TE, 3 levels for anterior/posterior, F7/F8, T3/T4, T5/T6). The ANOVAs included Sentence Type (4 levels, canonical SOV, non-canonical OSV, ambiguous SOV, ambiguous OSV), Anterior-Posterior factor (5 levels for PS Anova and 3 levels for TE Anova) and Hemisphere (2 levels). Based on the predictions made considering the previous experiments in the literature (Rösler et al., 1998; Matzke et al., 2002; Bornkessel et al. 2002, Schlesewsky et al., 2003; Hagiwara et al., 2007) we chose initially four time-windows of exploration: (i) 300-500 ms in the First determiner Phrase, where the LAN component was predicted for OSV sentences; (ii) 700-900 ms, at the verb position, based on the prediction that disambiguation should take place at the verb location in OSV ambiguous conditions; (iii and iv) 300-500 ms and 700-900 ms at the auxiliary position. These last two time-windows were included because visual inspection of the corresponding grand-averages showed that the effects elicited at the verb position continued until the end of the sentence (auxiliary word). The results of these ANOVAs are summarized at Table 4. All statistical tests comprised mean amplitudes for the different time windows specified in the corresponding contrast.

In order to decompose the interactions encountered in the omnibus analysis and to have a finer-grained analysis, further pairwise ANOVAs were conducted directly comparing (i) unambiguous sentences: SOV vs. OSV, (ii) ambiguous sentences, SOV vs. OSV and (iii) ambiguous vs. unambiguous sentences. These pairwise ANOVAs were carried out also at PS and TE locations including the corresponding Anterior/Posterior and Hemisphere factors. When the ANOVA was applied to midline locations (MD, Fz, Cz, Pz) only the Anterior/Posterior factor was included in the design.

For all statistical effects involving two or more degrees of freedom in the numerator, the Hynh-Feldt epsilon was used to correct for possible violations of the sphericity assumption. The exact *P*-value after the correction will be reported bellow. All ERP waveforms displayed in the corresponding figures were digitally filtered using a low-pass filter with a 12 Hz half-power cut off.

RESULTS

Behavioral data showed that participants performed the experiment very well. The comprehension task showed a mean percentage of correct responses of 91 % ± 7.8.

The results of the omnibus ANOVAs including all the experimental conditions are summarized at Table 4. There were significant interactions between sentence type, brain hemisphere and anterior/posterior localizations in each of the time-windows explored which indicated that the different types of sentences were processed differently. These interactions were therefore followed up by separate pairwise ANOVAs performed at PS, TE and MD locations and for all sentence's elements (including DP2) in order to study separately the effects of word order and ambiguity.

Unambiguous sentences

The direct comparison between SOV and OSV unambiguous sentences showed three main ERP effects (see Figure 3-5). At sentence initial position, an increased negativity between 300 and 500 ms distinguished Subjects (S) from Objects (O) (Figure 3). The onset of this component was approximately at 300 milliseconds after initial O sentences, and lasted until 500 milliseconds. This effect showed a wide spread distribution, being maximally at left frontal and extending to more central and posterior locations (left parietal and occipital), as the interactions between sentence type (ST) hemisphere (H) and anterior/posterior (AP) factors showed (300-500 ms, TE: Sentence type x Hemisphere x Anterior/Posterior, $F(2,46) = 8.68$, $P < 0.005$). The decomposition of this interaction in each Anterior/Posterior locations (frontal, central and posterior sides), showed that the crucial ST x H interaction was only significant at the frontal locations ($F(1,23) = 12.9$, $P < 0.002$; central locations, $F = 1.5$, $P < 0.23$; posterior ones, $F < 1$). This interaction reflects that while at F7 locations a negative increase is observed for OSV sentences, this effect showed inverted polarity at the right frontal locations, being this condition more positive (F8 location).

At the second DP position (see Figure 4), a finer-grained pairwise analysis revealed a second left negativity between 400 and 550 milliseconds distinguishing SOV and OSV sentences (400-550 ms, TE: Sentence type x Hemisphere, $F(1,23) = 5.01$, $P < 0.035$). The decomposition of this interaction, showed that Sentence type effect was only significant at the left hemisphere ($F(1,23) = 6.54$, $P < 0.017$), but not at the right hemisphere ($F < 1$). In the topographical maps of Figure 4 it is shown that distribution of this component is left lateralized extending to temporal locations.

Finally, at the verb position (Figure 5), a P600 component in the OSV sentences signaled the increased syntactic computation required to process this condition. This P600 effect was observed between 700 and 900 milliseconds after the appearance of the verb (PS: Sentence type x Anterior/Posterior, $F(4,92) = 7.92$, $P < 0.0014$). The decomposition of this interaction showed that Sentence type effect was only significant at parietal locations (P3-P4, $F(1,23) = 4.7$, $P < 0.05$). This effect reflects the standard posterior distribution of the P600 component (see topographical maps at Figure 5).

Ambiguous Sentences

In our experiment, syntactically fully-ambiguous sentences (examples 9c and 9d in Table 1) were disambiguated at the verb position. At this position, participants relied in their world knowledge in order to choose a plausible interpretation for the ambiguous sentence. Recall that the Determiner Phrases in these conditions carried the *-ak* ending, which are ambiguous and could be interpreted as a singular Subject or as a plural Object (see Table 1, examples 9c and 9d). Participants could not establish the syntactic function of the constituents (S or O) even upon reading the verb, where they had to resort to world-knowledge to choose one of the two possible readings of the sentence: as shown in example 9c, for instance, the sentence "otso-ak ardi-ak jan ditu" can mean either *the wolf has eaten the sheep (plural)* or *the sheep has eaten the wolves*, with world-knowledge clearly favoring the former interpretation over the latter. Thus, as we expected, the ERPs did not show any difference between DPs in these fully ambiguous conditions.

As it is shown in Figures 6 and 7, at sentence final position we observed a long lasting negativity with an onset approximately at 500 ms after the appearance of the verb (time-window, 500-1000 ms, TE: Sentence type x Anterior/Posterior, $F(2,46) = 5.72$, $P < 0.006$; PS: Sentence type x Hemisphere, $F(4,96) = 4.33$, $P < 0.008$). At temporal locations (TE) the decomposition of the Sentence type x Anterior/Posterior interaction showed that the effect of ST was significant only at frontal locations (F7-F8; $F(1,23) = 5.6$, $P < 0.002$; lateral central and posterior ones, in both cases, $F < 1$). At parasagittal sides, the decomposition of the interaction Sentence type x Hemisphere, showed that ST was only significant at the right hemisphere ($F(1,23) = 4.94$, $P < 0.036$), but not at the left hemisphere ($F(1,23) = 3.11$, $P < 0.0913$). In sum, this negative component was right frontally distributed in verb position and auxiliary positions (see topographical maps at Figure 6).

At Figure 7, we also plotted the unambiguous SOV condition for the sake of comparison with the ambiguous SOV and OSV conditions at verb position. Notice that an increase of positivity (P600) is observed for ambiguous SOV conditions when compared to the unambiguous ones. A pairwise comparison between both SOV conditions (unambiguous vs, ambiguous) (time window 700-900 ms after verb onset) showed a significant difference between them in the P600 component (e.g., at Po2, Pz, P3 locations, in all cases $F(1,23) > 4.8$, $P < 0.03$).

Discussion

In the last ERP experiment we investigated the electrophysiological correlates of canonical (SOV) and non-canonical word orders (OSV) in ambiguous and unambiguous syntactic contexts. The main results confirmed the evidences obtained in the previous behavioral experiments with regard to the higher processing cost derived from the higher syntactic complexity of non-canonical Object-first sentences (Experiment 1), and with regard to the existence of default subject-first processing mechanism observed in Basque, which is also applied in the case of fully ambiguous sentences (Experiment 2). In the following sections we discuss the main ERP findings obtained and their interpretation considering the different positions at the sentence in which they were elicited.

Sentence initial position

Regarding the initial position in the unambiguous sentences (SOV and OSV), the ERPs showed an increased negativity at object position (Figure 3), being this effect visible at left anterior locations and inverting the polarity at the right frontal locations (increased positivity for object particles). Because the spread distribution of this component and the time-window in which it is observed (300-500) it is not possible to sort out if this component is reflecting the overlap between a left anterior negativity and a N400 components, or it is exclusively a N400 effect. However, as in the case of object-first sentences in German (Rösler et al., 1998, Matzke et al., 2002, Felser et al., 2003), we can conclude that syntactically fronted/displaced constituents generate electrophysiological differences because the structure of the sentences with displaced constituents is more complex and requires further processing operations.

In Japanese, when the object appeared before the subject, it did not generate any significant ERP component at a sentence initial object position comparing with sentence initial subjects (Hagiwara et al., 2007). These results are in accordance with some behavioral studies using self-paced reading in Japanese where no differences were encountered between SOV and OSV word orders (Yamashita, 1997; Tamaoka et al., 2003), but they do not converge with some other studies using self-paced reading (Miyamoto and Takahashi, 2002) or combining self-paced reading and eye-tracking measures (Mazuka et al., 2002).

In the Japanese ERP experiment (Hagiwara et al., 2007), when the object was extracted out of the embedded clause so that the subject of the main clause appeared in between the object and the subject of the embedded clause, the LAN component was observed, presumably due to the increase in syntactic complexity. Even if the context where the LAN component appeared in Japanese does not seem exactly the same context as the one in which our negative LAN/N400 component was observed, its topographical distribution is somehow similar, showing a left-frontal-occipital distribution (see figure 3). Hagiwara et al. (2007) suggest that this distribution is related to the fact that in Japanese Case markers are attached at the end of words (in Japanese, ken-*ga*; as in Basque, emakume-*ak*), and therefore the parser might first have access to the noun information and then (or simultaneously) to grammatical information. In contrast, in languages like German, grammatical Case is marked before the words they are related to (e.g., *der* Mann) and thus the parser can initially begin to process syntactic information. Hagiwara et al. (2007) reasoned that the occipito-temporal distribution of the LAN component might reflect the differential initial access to the semantic information in both types of languages, which is assumed to be processed at the left-temporal lobe. In this regard, the similarity of the scalp distribution of our negative component with the LAN component reported by Hagiwara et al. (2007) provides further evidence for the relevance of the head-parameter, reflecting the modulation of parametric differences between head initial (English, German, etc.) and head final languages (Japanese, Basque, etc.).

Second position words

At second position, in all conditions, all constituents had the same *-ak* ending (examples 9 in Table 1). In the unambiguous conditions, the preceding word (S in the SOV condition and O in the OSV condition) determined whether the second word was plural O or singular S. When comparing the second position of the unambiguous condition, a left negative component (temporal-posterior scalp distribution, see Figure 4) was observed in the case of S. In German, comparing canonical and non-canonical word orders, Matzke et al. (2002) observed a LAN-like component at S position when this constituent was processed after O. Unlike our materials, sentences in Matzke et al. (2002) had the Verb located in between S and O (SVO and OVS word orders). The authors interpreted this second LAN component at S position of OVS sentences as an indicator of the retrieval of verbal material in a non-canonical position. But using Verb-final sentences, Rösler et al. (1998) also found a LAN-like component in second Object position (non-canonical in German embedded clauses: S-*O*-IO-V) compared to canonical Indirect Objects (canonical in German embedded clauses: S-*IO*-O-V). Following the ERP literature about the sentence initial LAN component (Kluender and Kutas, 1993; King and Kutas, 1995; Müller et al., 1997; Münte et al., 1998), Rösler et al. (1998) suggested that this LAN-effect indicates an extra load on working memory, until the corresponding noun phrase can be interpreted in its canonical position. We can extend this interpretation to the negative component showed in second argument position of non-canonical unambiguous OSV condition in Basque.

Regarding the left temporal-posterior distribution of our LAN-like effect, it is important to remark the variable topography of the LAN component in the literature. Some studies have reported a clear left frontal distribution (Neville et al., 1991; King and Kutas, 1995; Osterhout & Holcomb, 1992; Rösler et al., 1993), other studies reported a frontocentral distribution (Coulson et al., 1998; Münte and Heinze, 1994; Münte et al., 1997), bilateral (e.g., Friederici et al., 1999), in the right hemisphere (Osterhout and Nicol, 1999, exp. 2; Linares et al., 2006) and still others report a left posterior temporal distribution (Rodriguez-Fornells et al., 2001; Newman et al., 2007). The scalp distribution of all these anterior negativities could depend on the superimposition of other components, the utilization of different types of stimuli, different languages, indirect tasks or individual differences. For example, in Japanese, the S of OSV (compared to the O of SOV) showed a frontal negativity, but in the right hemisphere (Hagiwara et al., 2007). According to the authors, this component was related to the fact that Japanese participants expected a transitive verb after the O, and not the S. Moreover, in more complex sentences with long displacement of the O (O-Top-S-V) a left lateralized P600 component was found at S position. The authors interpreted the left positivity as reflecting the cost of the structural integration of long-displaced O in S position.

Our left temporal-negative component in sentence second position of unambiguous conditions (examples 9a and 9b in Table 1) converges with the interpretation given by Rösler et al. (1998). That is, the LAN component expresses the storage of displaced elements in working memory (Gibson, 1998; Felser et al., 2003) or, alternatively, it expresses that the Subjects and the Objects are processed differently regardless to their position.

Sentence final position

The Verb heads the Verb Phrase, and therefore, once the verb is processed, the structure of the Verb Phrase is established. When processing a canonical structure, no extra processing cost is required at V position. But when a Verb Phrase involves elements displaced from their canonical positions, then the processing cost increases at verb position in order to recover the required syntactic information. Our electrophysiological results revealed a P600 component at the V position in the unambiguous OSV condition (example 9b in Table 1) signaling the higher syntactic complexity of these non-canonical word orders. This interpretation converges with the hypothesis that the P600 component is related to an increase in processing load at V position, for example when comparing Subject vs. Object relative sentences and questions (Kluender and

Kutas, 1993; King and Kutas, 1995; Müller et al., 1997; Kluender and Münte, 1998; Phillips et al., 2001; Fiebach et al., 2002; Felser et al., 2003).

Regarding ambiguous sentences (see examples at 9c and 9d in Table 1), disambiguation occurred at sentence final position, driven by world knowledge. When participants read the verb, they chose the most plausible interpretation according to their knowledge of the world (semantic processing). In the case of the ambiguous SOV condition (example 9c in Table 1) the interpretation of the sentence did not require any syntactic reanalysis. But in the other case (example 9d in Table 1), once the two ambiguous words were read, participants applied a default processing mechanism that drove them to start with a Subject-then-Object interpretation. However, when they read the verb, the most plausible interpretation required them to change from the canonical S-O preference to the non-canonical O-S order.

In previous studies combining syntactic and semantic violations, ERPs showed N400 and P600 components (Hahne and Friederici 2002; Hagoort, 2003) or an early negativity and P600 (Hahne and Friederici, 2002). However, in our experiment, ERPs showed a large right frontal negativity that stayed significant to the end of the sentence. As far as we know, there are no experimental investigations where a semantic disambiguation causes a syntactic reanalysis. A previous ERP experiment with double violations (semantic and syntactic) obtained a right anterior negativity when compared to the single violation conditions (Osterhout and Nicol, 1999). In a similar vein, Linares et al. (2006) obtained a similar right anterior negativity in a double violation condition in which an incorrect suffix and stem violation in irregular Spanish verbs were presented.

These anterior negative components could be related to the amount of effort required for processing double violations irrespective of whether they are semantic, syntactic, or morphological violations. Bearing this in mind, we hypothesize that in the case of a syntactic reanalysis forced by world-knowledge, this frontal negative component observed at verb position could reflect the increased level of effort or cognitive control required to integrate the meaning of the sentence in the hypothesized word order. Because cognitive control recruits the frontal executive system (Braver and Ruge, 2006), this neural network might explain the bilateral frontal topography of the negative component observed at the disambiguation point.

Besides, the present slow negativity potential shift component at the end of the sentence is also reminiscent of previous slow sentence components that have been reported in previous studies (King and Kutas, 1995; Münte et al., 1998; Muller et al., 1998; Vos et al., 2001). The functional role attributed to this type of negative components (some of them frontal or left-frontally distributed) is the increased load in working memory required to syntactic processing. For example, Münte et al. (1998) encountered a slow negativity that was present for sentences beginning with the word "before" relative to those beginning with the word "after". These authors argued that increased semantic load was present in "before" type sentences because the temporal order of events had to be rearranged when compared to the canonical temporal order in "after" type sentences. Because of that, the first part of the sentence had to be maintained in working memory until it could be integrated into the correct order. The presence of this frontal negative component in our study at the end of the OSV ambiguous sentences might similarly reflect the memory load associated to syntactic reanalysis. In this type of sentences the participant might need to rearrange the constituents because the initial default SOV interpretation.

Conclusions

Our study shows that certain previous findings in ERPs language-processing studies observed in head initial, nominative and relatively fixed word order languages, are also observed in a head final, ergative, free word order and highly inflected and ergative language like Basque. This strongly suggests that these findings signal universal processing mechanisms, independent on parametric specifications of the grammar. We have also found strong evidence for a canonical word order processing advantage: even though practically all phrase-permutations are possible in Basque, SOV word order, argued by linguists to be canonical is processed significantly faster and with greater ease. Ambiguity resolution further supports this conclusion, because it favors a canonical SOV word order processing. Our results therefore provide evidence for the overall validity of neuroimaging measures of syntactic complexity across language types, and moreover they provide independent supporting evidence that even languages displaying free word order have an underlying canonical order, corresponding to the least complex sentence-structure.

Word order in human languages reflects syntactic structure, so that some sequences of phrases in a sentence correspond to simpler structures whereas other sequences correspond to more complex structures. The word order resulting from the simplest sentence structure is referred to as *canonical* in linguistics, and all others are usually referred to as *derived* word orders, because they involve further syntactic computation. We have shown that canonical order is processed faster and with greater ease, presumably because it involves a simpler computation than derived ones. We have shown that even in highly inflected and free-word order languages like Basque, there is clear behavioral and electrophysiological evidence for a canonical, easier and faster to process sequence of phrases in a sentence. Thus, this appears to be a universal design property of language, despite differences on case marking, verb agreement and other variable specifications of a given grammar.

Acknowledgements

This research has been supported by research grants of the European Science Foundation (BFF2002-20379-E), the Spanish Ministry of Education and Science (CSD2007-00012, SEJ2007-60751/PSIC) and the Basque Government (GIU06/52) to Itziar Laka and Kepa Erdozia, a pre-doctoral grant from the University of the Basque Country to Kepa Erdozia, the Spanish Government (SEJ2005-06067/PSIC) to Antoni Rodriguez-Fornells and a predoctoral grant from the Spanish Government to Anna Mestres-Missé. We want to thank Ruth de Diego Balaguer for her comments on a previous draft of this manuscript, Elisabeth Gilboy for her comments and helping us in the preparation of the self-paced reading experiments and Toni Cunillera for technical assistance.

References

Arregi, K. (2002). *Focus on Basque movements*. Ph.D. dissertation. MIT.

Assadollahi, R., and Pulvermüller, F. (2001). Neuromagnetic evidence for early access to cognitive representations. *Neuroreport*, 12(2), 207-213.

Bahlmann J., Rodriguez-Fornells A., Rotte M., Münte T. F. (2007). An fMRI study of canonical and noncanonical word order in German. *Human Brain Mapping*.

Baker, M. C. (2001). *Atoms of Language: The mind's hidden rules of grammar*. Basic Books: NY.

Bates, E., Friederici, A., and Juarez, L. (1988). On the preservation of word order in aphasia. *Brain and Language*, 33(2), 323-364.

Bates, E., Devescovi, A., Wulfeck, B. (2001). Psycholinguistics: a cross-language prespective. Annual Review of Psychology. 52, 369-396.

Ben-Shachar, M., Palti, D., Grodzinsky, Y. (2004). Neural correlates of syntactic movement: converging evidence from two fMRI experiments. *NeuroImage*, 21, 1320-1336.

Bever, T.G. (1970). The cognitive basis for linguistic structures. In J.R. Hayes (ed.). *Cognition and the development of language*. Wiley: NY., 277-360.

Bornkessel, I., Schlesewsky, M., Friederici, A. (2002). Grammar overrides frequency: evidence from online processing of flexible word order. *Cognition*, 85, B21-B30.

Bornkessel, I., Fiebach, C. J., Friederici, A. D. (2004). On the cost of syntactic ambiguity in human language comprehension: an individual differences approach. *Cognitive Brain Research*, 21, 11-21.

Bornkessel, I., Zysset, S. Friederici, A., von Cramon, Y., and Schlesewsky, M. (2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. *NeuroImage*, 26, 221-233.

Bornkessel, I., Schlesewsky, M. (2006). The extended argument dependency model: a neurocognitive approach to sentence comprehension across languages. *Psychological Review*, 113(4), 787-821.

Braver, T.S., and Ruge, H. (2006). Functional Neuroimaging of Executive Functions. In R. Cabeza and A. Kingstone (Eds.), *Handbook of Functional Neuroimaging of Cognition* (pp. 307-348). Cambridge, MIT Press.

Brown, C. M., Hagoort, P., and ter Keurs, M. (1999). Electrophysiological signatures of visual lexical processing: open- and closed-class words. *Journal of Cognitive Neuroscience*, 11, 261-281.

Chomsky, N. (1981). *Lectures on Government and Binding*. Foris, Dordrecht.

Coulson, S., King, J. W., and Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes*, 13, 21-58.

Davis, C.J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, 37, 65-70.

De Rijk, R. (1969). Is Baque an S.O.V. language? *Fontes Linguae Vasconum* 1-3, 319-351. Reedited by De Rijk, Rudolf P. G., 1998, *De Lingua Vasconum: Selected Writings*. Supplements of the *Anuario del Seminario de Filología Vasca "Julio de Urquijo"* nº XLLIII.

De Rijk, R. (2007). *Standard Basque, A Progressive Grammar*, MIT Press, Cambridge MA.

Dixon, R. (1994). *Ergativity*. Cambridge Studies in Linguistics 69. Cambridge University Press, UK.

Elordieta, A. (2001). *Verb Movement and Constituent Permutation in Basque*. LOT Dissertation Series: Utrecht.

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

Fiebach, C. J., Schlesewsky, M., and Friederici, A. (2002). Separating syntactic integration cost during parsing: the processing of German WH-questions. *Journal of Memory and Language*, 47, 250-272.

Fiebach, C.J., Schleswsky, M., Lohmann, G., von Cramon, Y., and Friederici, A. (2005). Revisiting the role of Broca's area in sentence processing: syntactic integration versus syntactic working memory. *Human Brain Mapping*, 24, 79-91.

Frazier, L., and Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6, 291-325.

Friederici, A. D., Steinhauer, K., and Frisch, S. (1999). Lexical integration: sequential effects of syntactic and semantic information. *Memory and Cognition*, 27, 438-453.

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

Friederici, A. D. and Kotz, S. A. (2003). The brain basis of syntactic processes: functional imaging and lesion studies. *NeuroImage*, Volume 20, S8-S17.

Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I. and Anwander, A. (2006). The brain differenciates human an non-human grammars: Functional localization and structural connectivity. *Proceeding of the National Academy of Sciences*, 14, vol. 103, 2458-2463.

Frisch, S., Schlesewsky, M., Saddy, D., and Alpermann, A. (2002). The P600 as an indicator of syntactic ambiguity. *Cognition*, 85, B83-B92.

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

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (ed.). *Universals of Language*. Cambridge, MA: MIT Press. Second printing, first paperback edition. 1966. 73-113.

Hagiwara, H., Takahiro, S., Masami, I., and Imanaka, K. (2007). A topographical study on the event-related potential correlates of scrambled word order in Japanese complex sentences. *Journal of Cognitive Neuroscience*, 19, 175-193

Hagoort, P. (2003). Interplay between Syntax and Semantics during Sentence Comprehension: ERP Effects of Combining Syntactic and Semantic Violations. *Journal of Cognitive Neuroscience*, 15, 883-899.

Hahne, A. and Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research*, 13, 339-356.

Hauk, O. and Pulvermüller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology*, 115, 1090-1103.

Hualde, J. I., and Ortiz de Urbina, J. (eds.). (2003). *A Grammar of Basque*. Mouton de Gruyter: Berlin.

Just, M. A., Carpenter, P. A., and Woolley, J. D. (1982). Paradigms and Processes in Reading Comprehension. *Journal of Experimental Psychology: General*, vol. 111(2), 228-238.

Kaan, E., (1997). *Processing subject-object ambiguities in Dutch.* PhD. Thesis: University of Groningen (*Groningen Dissertations in Linguistics 20*).

Kaan, E., and Swaab, T. Y. (2002) The brain circuitry of syntactic comprehension. *Trends in Cognitive Sciences,* Volume 6, Issue 8, 350-356.

Kennedy, A., and Murray, W. S. (1984). Reading without eye movements. In A. G. Gale and F. Johnson (eds.). 1984. *Theoretical and Applied Aspects of Eye Movement Research*. North-Holland.

King, J., and Kutas, M. (1995). Who did what when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7, 376-395.

Kluender, R., and Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, 5, 196-214.

Kluender, R., and Münte, T. (1998). Subject/object asymmetries: ERPs to grammatical and ungrammatical wh-questions. Poster presented at the 11th annual CUNY conference on human sentence processing. Rutgers University.

Kutas, M., Van Petten, C., and Kluender, P. (2006) Psycholinguistics electrified II. In M. Traxler and M. A. Gernsbacher (eds.), *Handbook of psycholinguistics* (2nd ed.). Amsterdam: Elseiver.

Laka, I. (1996). *A Brief Grammar of Euskara, the Basque Language.* University of the Basque Country. Internet document available at http://www.ehu.es/grammar

Linares, R., Rodriguez-Fornells, A., and Clahsen, H. (2006). Stem allomorphy in the Spanish mental lexicon: Evidence from behavioral and ERP experiments. *Brain and Language*, 97, 110-120.

Mak, W., Vonk, W., and Schriefers, H. (2006). Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language*, 54, 466-490.

Matzke, M., Mai, H., Nager, W., Rüsseler, J., and Münte, T. (2002). The cost of freedom: An ERP-study of non-canonical sentences. *Clinical Neurophysiology*, 113, 844-852.

Mazuka, R., Itoh, K., and Kondo, T. (2002). Costs of scrambling in Japanese sentence processing. In M. Nakayama (ed.), *Sentence Processing in East Asian Languages*. Stanford: CSLI.

Miyamoto, E. T., and Takahashi, S. (2002). Sources of difficulty in the processing of scrambling in Japanese. In M. Nakayama (ed.), *Sentence Processing in East Asian Languages*. Stanford: CSLI.

Mohr, B., Pulvermüller, F., and Zaidel, E. (1994). Lexcial decision after left, right, and bilateral presentation of function words, content words, and non-words: evidence for interhemispheric interaction. *Neuropsychologia*, 32, 105-124.

Müller, H., King, J., and Kutas, M. (1997). Event-related potentials elicited by spoken relative clauses. *Cognitive Brain Research,* 5,193-203.

Münte, T. F., and Heinze, H. J. (1994). ERP negativities during syntactic processing of written words. In: H.J. Heinze, T.F. Münte & G.R. Mangun (Eds.), *Cognitive Electrophysiology*. Boston: Birkhäuser.

Münte, T. F., Matzke, M., and Johannes, S. (1997). Brain activity associated with syntactic incongruencies in words and pseudo-words. *Journal of Cognitive Neuroscience,* 9, 318-329.

Münte, T. F., Heinze, H. J., Matzke, M., Wieringa, B. M., and Johannes, S. (1998). Brain potentials and syntactic violations revisited: no evidence for specificity of the syntactic positive shift. *Neuropsychologia*, 36, 217-226.

Münte, T.F., Schiltz, K., Kutas, M. (1998). When temporal terms belie conceptual order. Nature, 395, 71-73.

Nakano, Y., Felser, C., and Clahsen, H., (2002). Antecedent priming at trace positions in Japanese long-distance scrambling. *Journal of Psycholinguistic Research*, Vol. 31, No.5, 531-571.

Neville, H. J., Nicol, J., Barss, A., Foster, K. I., and Garrett, M. F. (1991). Syntactically based sentence processing classes: evidence from eventrelated brain potentials. *Journal of Cognitive Neuroscience,* 3, 151-165.

Newman, A., Ullman, M.T., Pancheva, R., Waligura, D., and Neville, H. (2007). An ERP study of regular and irregular English past tense inflection. *NeuroImage,* 34, 435-345.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia,* 9 (1), 97-113.

Ortiz de Urbina, Jon. (1989). *Parameters in the Grammar of Basque*. Foris: Dordrecht.

Osterhout, L. and Holcomb, P. J. (1992). Event-Related Brain Potentials Elicited by Syntactic Anomaly. *Journal of Memory and Language,* 31, 785-806.

Osterhout, L., Bersick, M., and McKinnon, R. (1997). Brain potentials elicited by words: Word length and frequency predict the latency of an early negativity. *Biological Psychology*, 46, 143-168.

Osterhout, L., and Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes*, 14, 283-317.

Pallier, C., Dupoux, E., and Jeannin, X. (1997). EXPE: an expandable programming language for on-line psychological experiments. *Behavior Research Methods, Instruments, & Computers*, 29(3), 322-327.

Perea, M., Urkia, M., Davis, C. J., Agirre, A., Laseka, E., & Carreiras, M. (2006). E-Hitz: A word-frequency list and a program for deriving psycholinguistic statistics in an agglutinative language (Basque). *Behavior Research Methods*, 38(4), 610-615.

Phillips, C., Kazanina, N., Wong, K., and Ellis, R. (2001). ERP evidence on the time course of processing demands in wh-dependencies. Poster presented at Human Sentence Processing 14. CUNY.

Rösler, F., Putz, P., Friederici, A., and Hahne, A. (1993). Event-Related Brain Potentials While Encountering Semantic and Syntactic Constraint Violations. *Journal of Cognitive Neuroscience,* 5, 345-362.

Rösler, F., Pechmann, T., Streb, J., Röder, B., and Hennighausen, E. (1998). Parsing of sentences in a language with varying word order: word-by-word variations of processing demands are revealed by even-related potentials. *Journal of Memory and Language*, 38, 150-176

Schlesewsky, M., Fanselow, G., Kliegl, R., and Krems, J. (2000). Preferences for grammatical functions in the processing of locally ambiguous wh-questions in German. In B. Hemforth and L. Konieczny (Eds.), *German Sentence Processing*, 65-93. Dordrecht: Kluwer.

Schlesewsky, M., Bornkessel, I., and Frisch, S. (2003). The neurophysiological basis of word order variations in German. *Brain and Language*, 86, 116-128.

Schlesewsky, M, and Bornkessel, I. (2006). Context-sensitive neural responses to conflict resolution: Electrophysiological evidence from subject-object ambiguities in language comprehension. *Brain Research*, 1098, 139-152.

Schwartz, B.D. and Vikner, S. (1996). The verb always leaves IP in V2 clauses. In: A. Belletti and Rizzi, L. (eds.), *Parameters and Functional Heads*. Oxford: OUP, 11-62.

Sekerina, I. (2003). Scrambling processing: dependencies, complexity, and constraints. In S. Karimi (ed.). *Word Order and Scrambling*. Blackwell: UK. 301-324.

Tamaoka, K., Sakai, H., Kawahara, J., and Miyaoka, Y. (2003). The effects of phrase-length order and scrambling in the processing of visually presented japanese sentences. *Journal of Psycholinguistic Research*, vol. 32, num 4, 431-454.

Traxler, M., Morris, R., and Selly, R. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language*, 47, 69-90.

Van Petten, C., and Kutas, M. (1990). Interactions between sentences context and word frequency in event-related potentials. *Memory and Cognition*, 18(4), 380-393.

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

Weber-Fox, C., and Neville, H. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8, 231-256.

Weekes, B.S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology*, 50A(2), 439-456.

Yamashita, H. (1997). The effects of word-order and Case marking information on the processing of Japanese. *Journal of Psycholinguistic Research*, 26, 163-188.

Table 1. Experimental materials for Experiment 1, 2 and 3 (S: Subject, O: Object, V: Verb; SOV: Subject-Object-Verb; OSV: Object-Subject-Verb)

EXPERIMENT 1		
	CANONICAL SOV	NON-CANONICAL OSV
	7a. Emakume-ak gizon-a ikusi du Woman-the(S) man-the(O) seen has(V) *'The woman has seen the man'*	7b. Gizon-a emakume-ak ikusi du Man-the(O) woman-the(S) seen has(V) *'The woman has seen the man'*
EXPERIMENT 2		
	CANONICAL SOV	NON-CANONICAL OSV
	8a. Emakume-ek gizon-ak ikusi dituzte Women-the(S) men-the(O) seen have(V) *'The women have seen the men'*	8b. Gizon-ak emakume-ek ikusi dituzte Men-the(O) women-the(S) seen have(V) *'The women have seen the men'*
	FULLY AMBIGUOUS	
	8c. Gizon-ak emakume-ak ikusi ditu Man-the(S/O) woman-the(S/O) seen has(V) *'The man has seen the women'* or *'The woman has seen the men'*	
EXPERIMENT 3		
	CANONICAL SOV	NON-CANONICAL OSV
Unambiguous	9a. Otso-ek ardi-ak jan dituzte Wolf-the(S) sheep-the(O) eat have(V) *'The wolves have eaten the sheep'*	9b. Ardi-a otso-ak jan du Sheep-the(O) wolf-the(S) eat has(V) *'The wolf has eaten the sheep'*
Temporally ambiguous	9c. Otso-ak ardi-ak jan ditu Wolf-the(S/O)sheep-the(S/O) eat has(V) *'The wolf has eaten the sheep'*	9d. Ardi-ak otso-ak jan ditu Sheep-the(S/O) wolf-the(S/O) eat has(V) *'The wolf has eaten the sheep'*

Table 2. Mean self-paced reading times (± standard deviation) in Experiment 1

	DP (ms)	***DP (ms)***	***V (ms)***	***Aux (ms)***	***Total (ms)***	***% errors***	*CT (ms)*
SOV	847 ±12	989 ±17	1046 ± 18	873 ± 9	3755 ± 49	4.3	2457 ± 24
OSV	789 ±13	1114 ± 21	1241 ± 22	1007 ± 14	4151 ± 59	15.6	2698 ± 30
SOV-OSV	*	*	**	*	**	***	**

Notes: DP: Determiner Phrase (subject or object depending on the condition), V: Verb, Aux: Auxiliary, CT: Comprehension Task; SOV: Subject-Object-Verb word order, OSV: Object-Subject-Verb word order. SOV-OSV: statistical comparison of two conditions (*t*-test), $P < .05$ = *; $P < .01$ = **; $P < .001$ = ***

Table 3. Mean self-paced reading times ± standard deviation in Experiment 2

	DP (ms)	*DP (ms)*	*V (ms)*	*Aux (ms)*	*Total (ms)*	*% errors*
SOV	1018 ± 17	1269 ± 26	1069 ± 20	1166 ± 25	4523 ± 80	5.7
OSV	937 ± 15	1471 ± 27	1349 ± 28	1653 ± 49	5412 ± 100	8.7
AMB	940 ± 15	1210 ± 21	1132 ± 27	1351 ± 24	4635 ± 78	12.8
SOV-OSV	*	*	***	***	***	
SOV-AMB	**			*		***
OSV-AMB		***	*		**	

Notes: DP: Determiner Phrase (subject or object depending on the condition), V: Verb, Aux: Auxiliary, CT: Comprehension Task; SOV: Subject-Object-Verb word order, OSV: Object-Subject-Verb word order, AMB: fully ambiguous sequences, SOV-OSV, SOV-AMB, and OSV-AMB: statistical two-by-two comparison of the experimental conditions (*t*-test), $P < .05$ = * ; $P < .01$ = ** ; $P < .001$ = ***

Table 4. Main ERP statistical results of the general ANOVAs conducted at Parasagittal and Temporal locations, at different positions of the sentence and at different time windows.

		DP1		Verb		Auxiliary			
		300-500ms		700-900ms		300-500ms		700-900ms	
	d. f.	F =	*p*(HF)	F =	*p*(HF)	F =	*p*(HF)	F =	*p*(HF)
Parasagittal.									
ST	3, 69							7.83	**
ST x AP	12, 276			2.18	*			3.75	***
Temporal									
ST	3, 69							11.07	***
ST x AP	6, 138			3.63	**	2.37	*		
ST x H x AP	6, 138	2.34	*						

Notes: ST: Sentence type (4 levels); AP: Anterior-Posterior factor; H: Hemisphere; DP1: first Determiner Phrase; d. f. : degrees of freedom; $p < .05$ = *; $p < .01$ = ** ; $p < .001$ = ***

Figures:

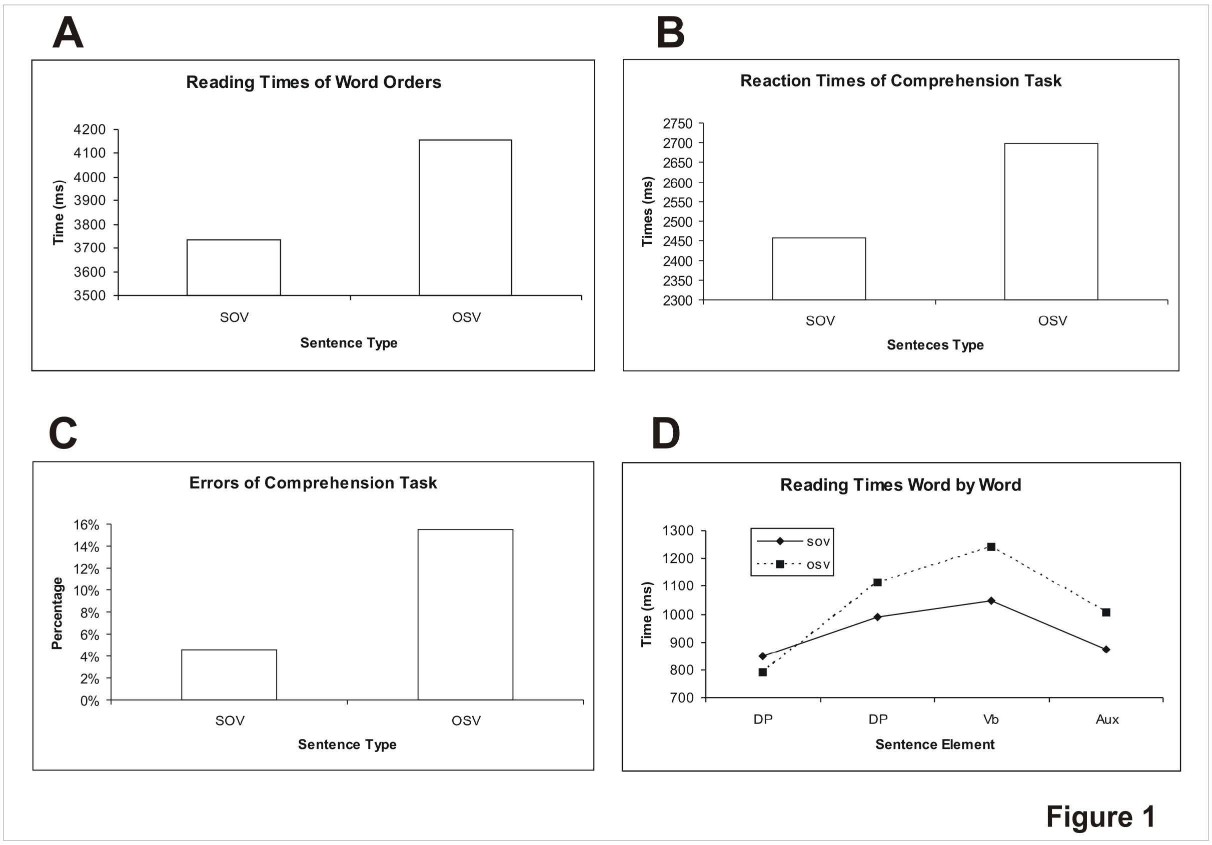

Figure 1. Results from the self-paced reading Experiment 1. A. Mean reading times to whole sentences in both experimental conditions (SOV = Subject-Object-Verb word order; OSV = Object-Subject-Verb word order). Notice that OSV condition required longer reading time than SOV condition. B. Mean reading time required by the participants to read and answer the questions correctly in both SOV and OSV conditions. C. Mean percentage of errors in the comprehension task for both experimental conditions. D. Word-by-word mean reading times in each condition. The y-axis features reaction times in milliseconds; in x-axis the different constituents of each sentence are depicted (DP = Determiner Phrase; Vb = Verb; Aux = Auxiliary verb). The DPs are the subjects and the objects of the transitive verbs. In the SOV condition the first DP corresponds to the subject of the sentence and the second one to the object, and vice versa in the OSV condition.

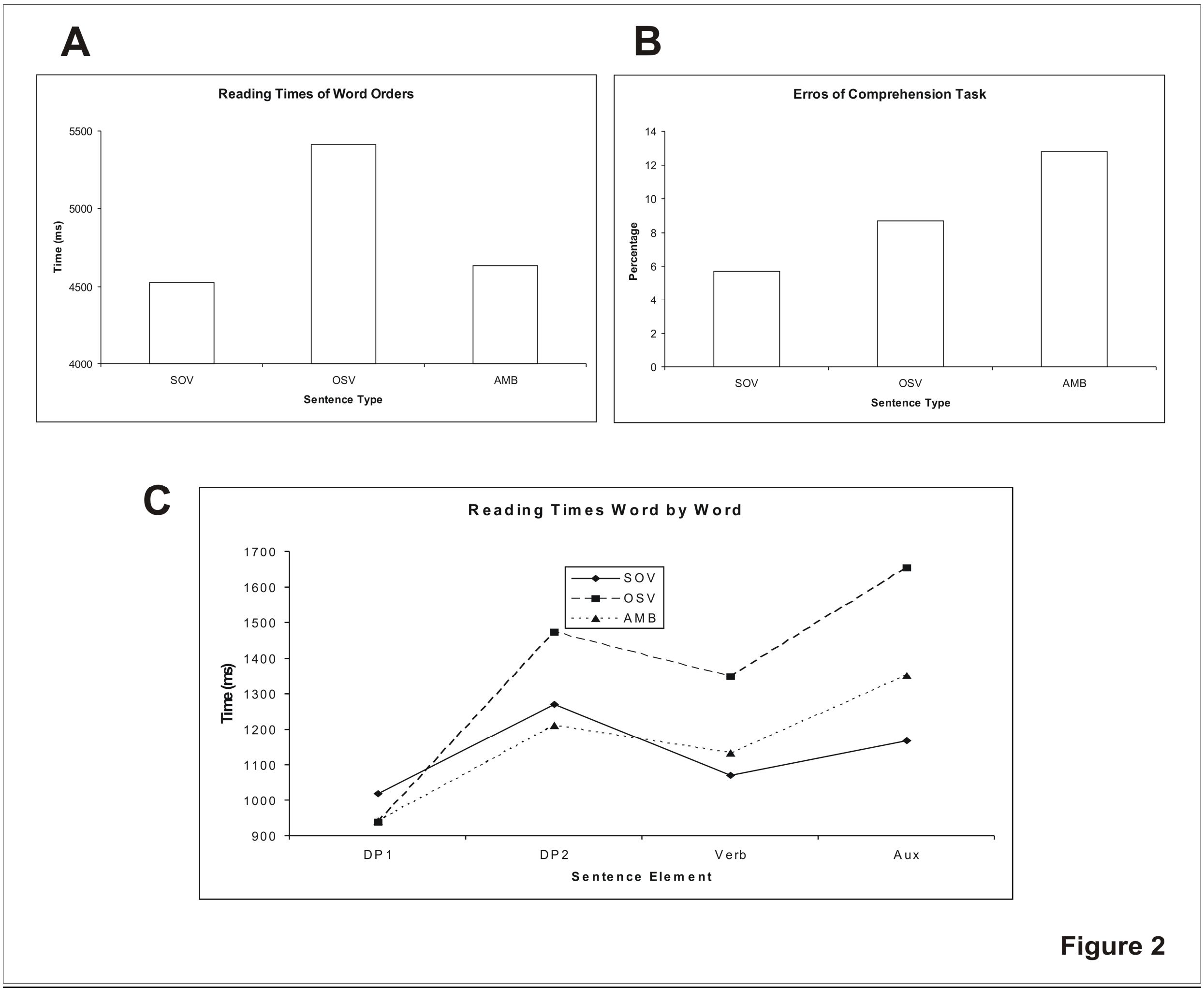

Figure 2. Results from self-paced reading Experiment 2. A. Mean reading times to whole sentences in both experimental conditions (SOV = Subject-Object-Verb word order; OSV = Object-Subject-Verb word order; AMB = fully ambiguous sequences). As in the first experiment, the OSV condition required longer reading times. B. Error rate in the comprehension task. C. Comparison of SOV, OSV and AMB conditions word-by-word. The y-axis features reaction times in milliseconds; in x-axis the different constituents of each sentence are depicted (DP = Determiner Phrase; Vb = Verb; Aux = Auxiliary verb). The DPs are the subject and the object of the transitive verb.

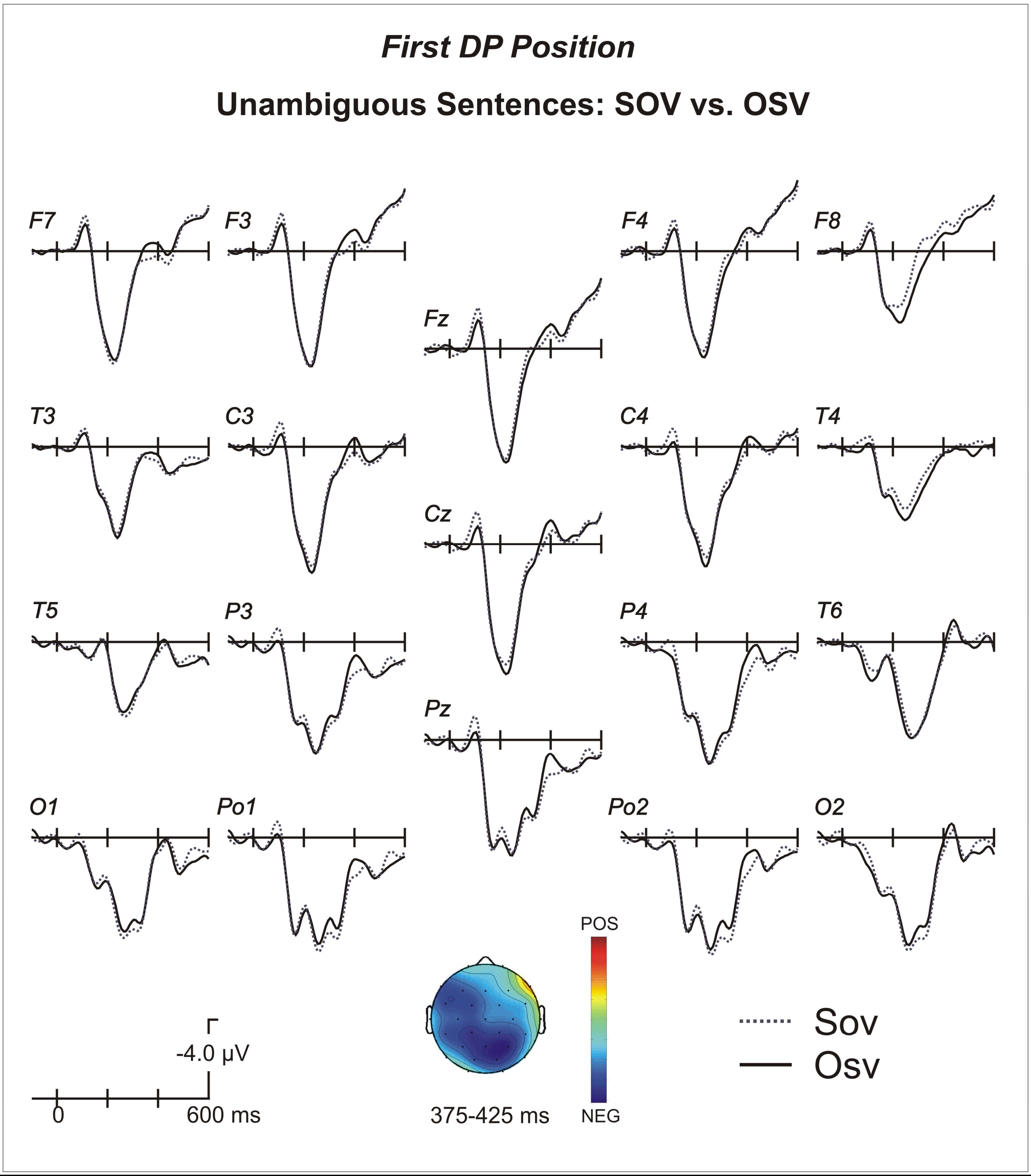

Figure 3. Results from the event related potentials Experiment 3. Unambiguous Subject-Object-Verb (SOV) and Object-Subject-Verb (OSV) sentences. The y-axis features the microvolt (the negative values are plotted up) and the x-axis features the time in milliseconds. Depicted the sentence first position (subject or object, Sov or Osv) at different scalp positions. An increased negativity was observed in object sentences in various electrode positions at 300-500 ms. The isovoltage topographical map showed the distribution of this negative component, obtained from

the subtraction of the object and the subject between 375 and 425 milliseconds after stimulus presentation.

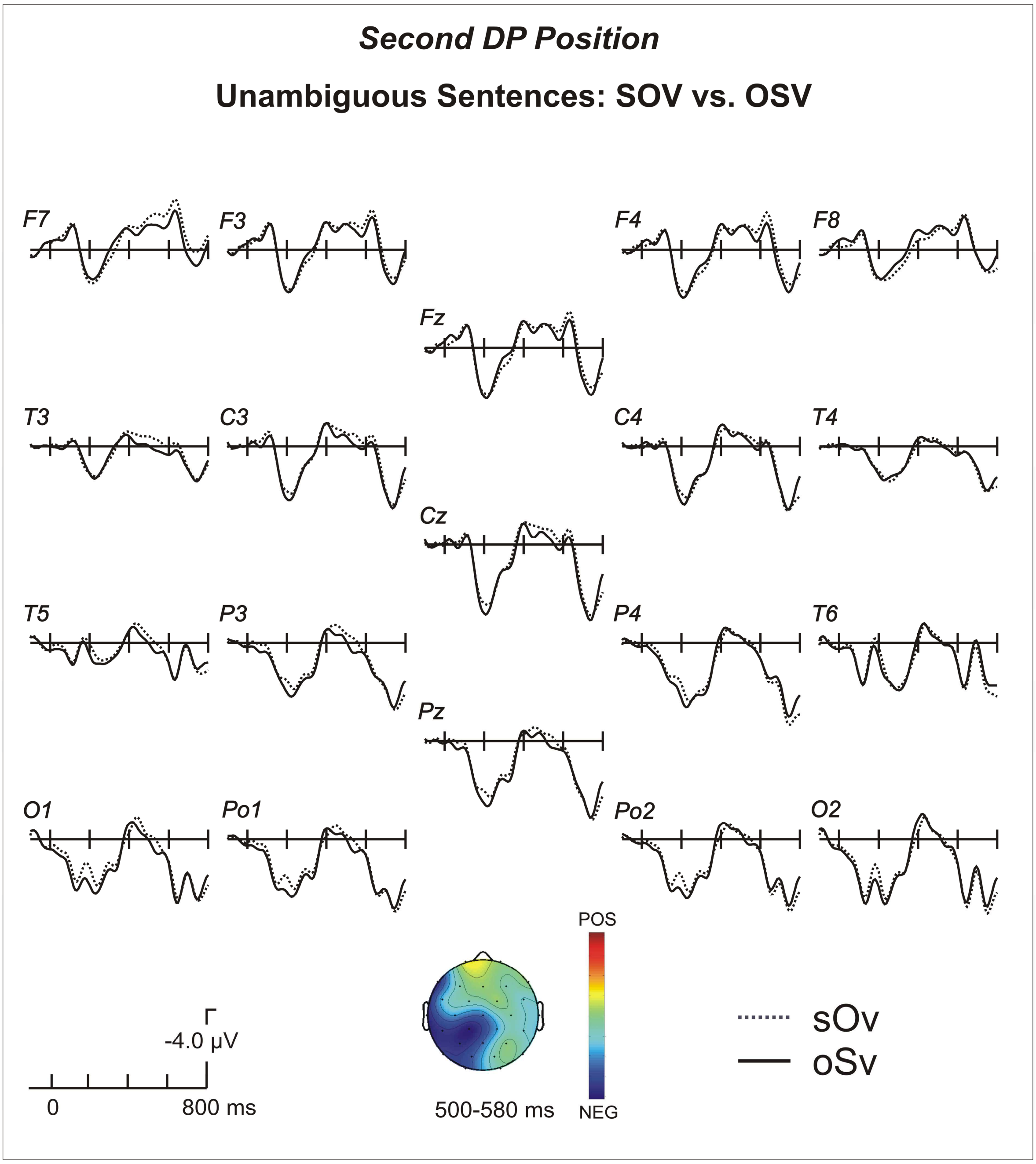

Figure 4. ERP comparison between unambiguous Subject-Object-Verb (SOV) and Object-Subject-Verb (OSV) sentences at sentence second DP position (subjects or objects). A second left negative component was observed in subject position of the non-canonical OSV condition

(oSv vs. sOv; notice that the capital letter signals the trigger of the averaging process). The topographical map showed that this component was localized in left middle temporal locations between 400 and 550 milliseconds.

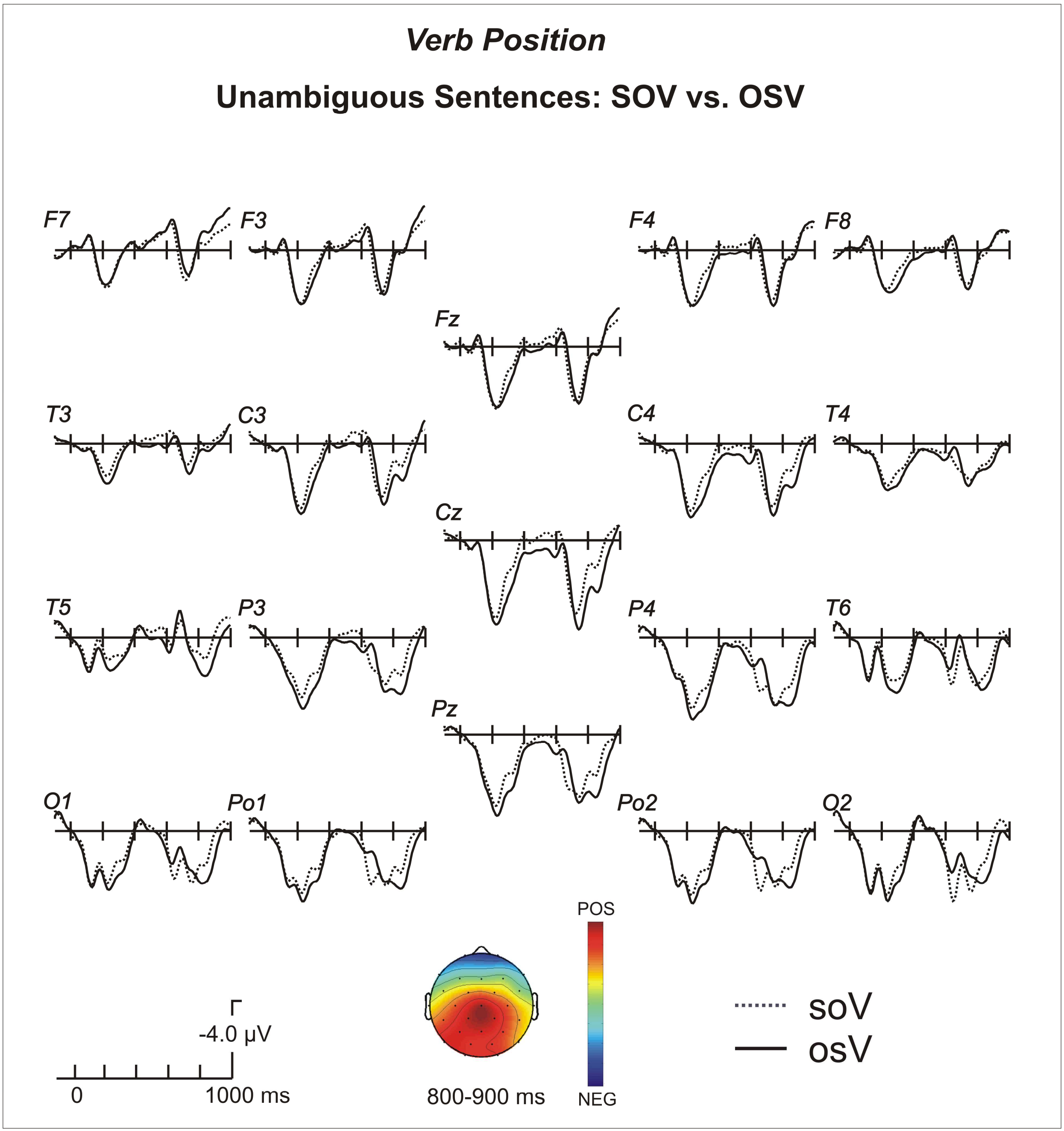

Figure 5. ERP comparison between unambiguous Subject-Object-Verb (SOV) and Object-Subject-Verb (OSV) sentences at the verb position. A clear P600 component was observed in

between 700 and 900 ms (comparing osV vs. soV sentences). The scalp distribution of this component showed the standard posterior parietal distribution (subtraction of verb position components of canonical SOV condition and non-canonical OSV conditions).

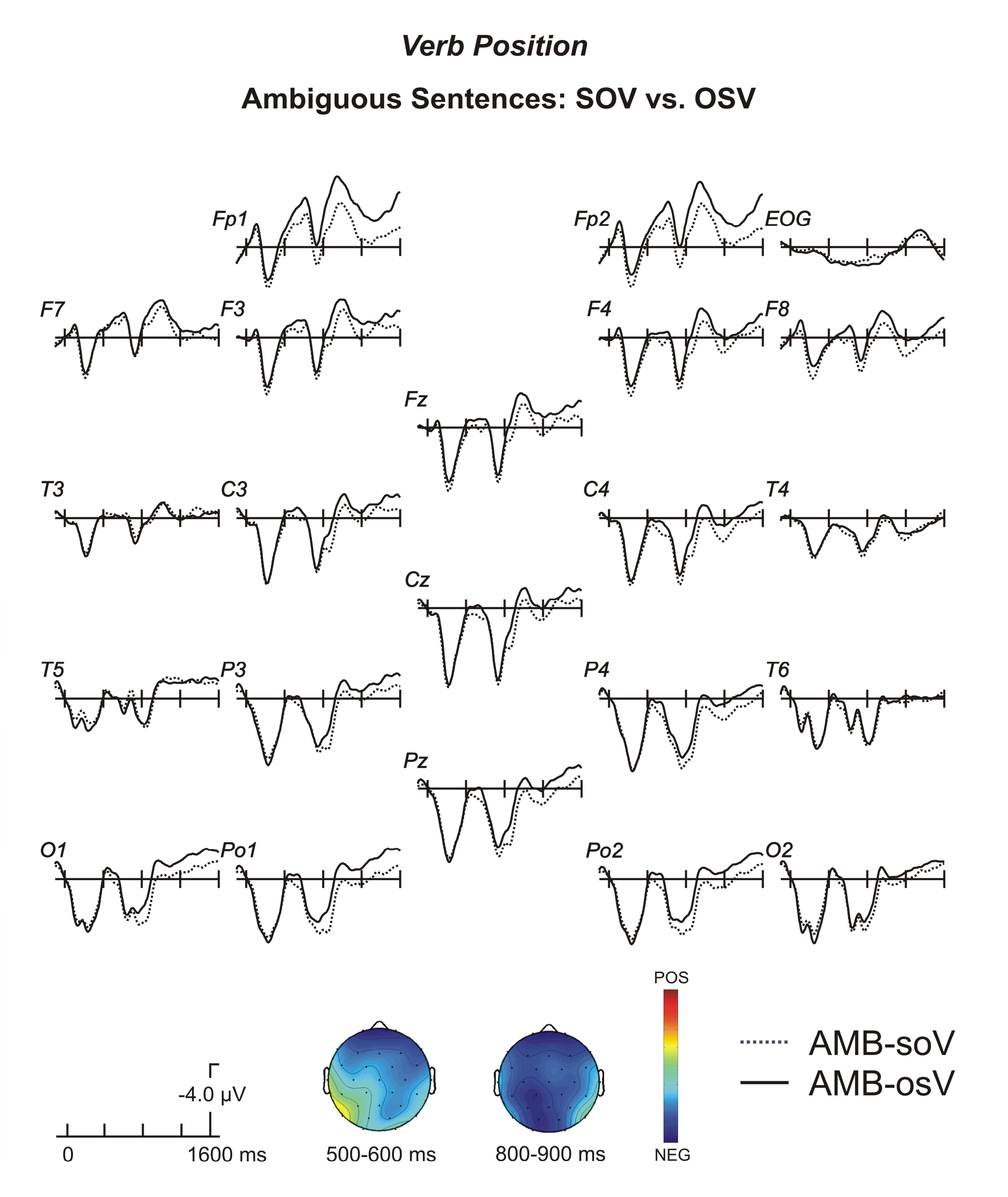

Figure 6. ERP comparison between ambiguous conditions at verb position (AMB-soV and AMB-osV). As it is shown, after 400 ms, a negative component was observed at frontal and posterior sides for the osV condition, specially at the period in which the auxiliary was processed. Notice that the time-window depicted in the electrodes is 1600 ms, which comprises the processing of the verb and the auxiliary.

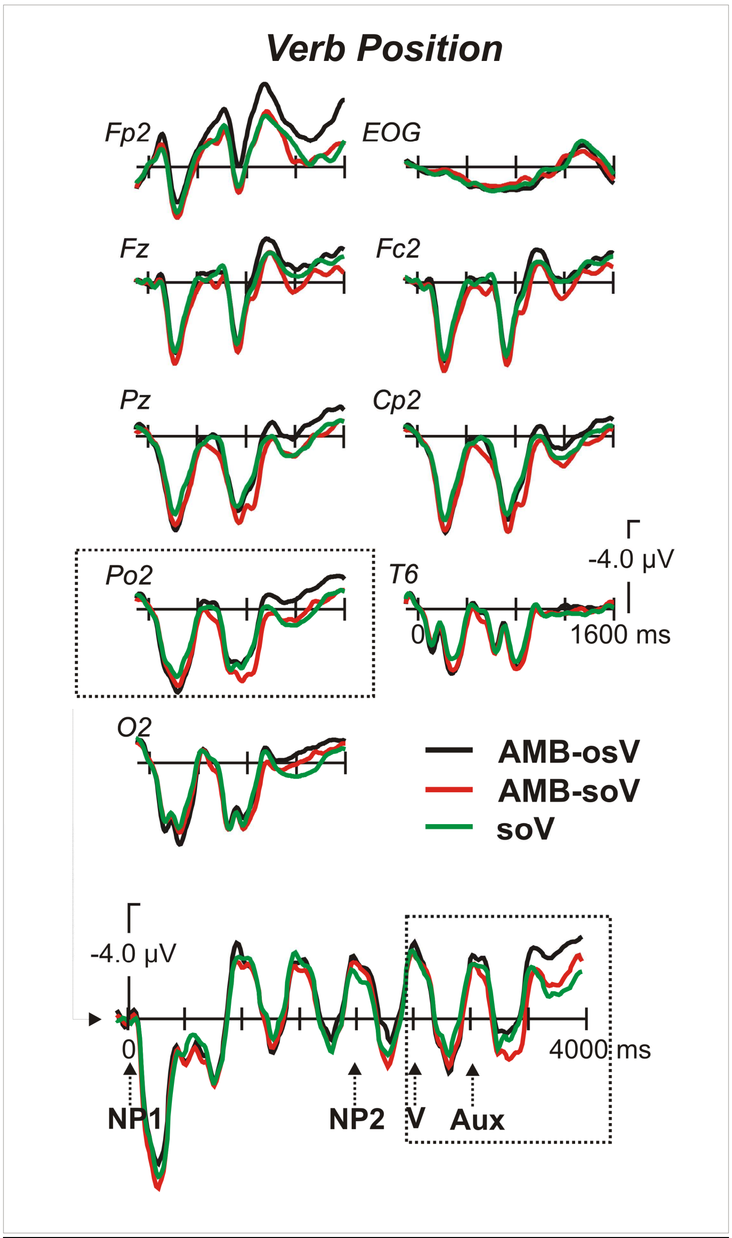

Figure 7. ERP results comparing ambiguous (AMB-osV and AMB-soV) and unambiguous (soV) conditions at verb position. The ambiguous SOV condition showed a P600 component

when compared to the unambiguous SOV condition. The down part of the Figure reflects the average of the sentences as a whole. We plotted 4 seconds epoch in Po2 electrode from the beginning to the end of both ambiguous conditions (AMB-SOV and AMB-OSV) and the unambiguous canonical (SOV).

Footnotes:

[1] The discarded 10 participants produce more errors in the comprehension task of non-canonical OSV condition (OSV = 67%; SOV =12.5%)

[2] In this experiment we were unable to compare the results of the comprehension task for the different conditions. The reason was that we introduced one extra word in the questions for the comprehension task of fully ambiguous sentences, which could directly affect the mean reaction time of this task. Recall that there were no errors in the comprehension task of fully ambiguous sentences, both answers (yes and no) were correct; however, because of experimental design reasons, we considered errors the OSV interpretation of fully ambiguous sentences.

[3] Taking together the 35 participants, OSV condition generated the most errors in the comprehension task (27.7%), followed by the ambiguous condition (15.4%) and by the SOV condition (6.9%).

[4] Assuming that the length of subjects and objects did not match in experiment 3, we investigated in which degree the ERP differences observed in the first DPs could be due to this differences. First of all, the length difference between Subjects and Objects was only 1.2 letters, which might suggest that the influence on the ERP morphology should not be too strong. In this regard, it has been shown that high-frequency words (which was the case in the present experiment) of different lengths were processed at a comparable speed (i.e., no word length effect), whereas low-frequency words showed word length effects (Weekes, 1997; see also Mohr et al., 1994). Thus, based on the existing psycholinguistic literature, the expected effect of a 1.2 letter difference should be small, specially compared to other effects like word frequency, which in the present case was well matched. The ERP literature about the effects of word length is less clear and more inconsistent (see Osterhout et al., 1997; Van Petten and Kutas, 1990; Brown et al., 1999). For example, Osterhout et al. (1997) reported a correlation of the latency of an early negativity and the mean normative frequency and mean length of the words regardless of word class. Other authors have described very early word length effects (see Assadollahi and Pulvermüller, 2001; Hauk and Pulvermüller, 2004). Interestingly, one of the clearest studies in which word length was investigated (Exp 2, from Van Petten and Kutas, 1990), the authors reported two effects related to the increase in word length: (i) a negative enhancement, peaking at 200 ms at parietal midline location (quantified as 150-225 ms time window at posterior sides), and (ii) a symmetrical broad positivity which appeared after 250 ms and continued for several hundred ms (mean voltage value in between 250-600 ms). However, the last effect appeared only in the comparison between the longest words (8 letters) with the shorter bins (3, 4-5, 6-7 letters). No differences in the second component were observed within the last bins. Following this study, we measure both components at midline locations (Fz, Cz, Pz) (Van Petten and Kutas, 1990). The first component (with a peak at 200 ms) showed a small increase in negativity for the Subjects vs. Objects, which in fact it should be expected considering van Petten and Kutas (1990). However, the corresponding ANOVA (at midline location, Fz, Cz, Pz) showed that this difference was not significant (Condition, $F(1,23) < 0.5$). No significant interaction between Condition and Electrode was encountered. Regarding the second component observed in Van Petten and Kutas (1990), the corresponding ANOVAs (250-600 ms) also showed no effect of Condition ($F(1,23) < 1$) or the corresponding interaction between Condition x Electrode ($F < 1$). Considering these analyses we believe that the effects reported in the present study might not be influenced by the length differences between objects and subjects. However, it is not completely possible to rule out this possibility.