Cognition 115 (2010) 79–92

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Subject relative clauses are not universally easier to process: Evidence from Basque

Manuel Carreiras [a,b,c,*], Jon Andoni Duñabeitia [a], Marta Vergara [d], Irene de la Cruz-Pavía [e], Itziar Laka [e]

[a] Basque Center on Cognition, Brain and Language, Donostia-San Sebastián, Spain
[b] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
[c] Departamento de Filología Vasca, Universidad del País Vasco (UPV-EHU), Spain
[d] University of California at Davis, CA, USA
[e] Departamento de Lingüística y Estudios Vascos, Universidad del País Vasco (UPV-EHU), Spain

ARTICLE INFO

Article history:
Received 1 June 2009
Revised 20 November 2009
Accepted 24 November 2009

Keywords:
Parsing
Relative clauses
Ergativity

ABSTRACT

Studies from many languages consistently report that subject relative clauses (SR) are easier to process than object relatives (OR). However, Hsiao and Gibson (2003) report an OR preference for Chinese, a finding that has been contested. Here we report faster OR versus SR processing in Basque, an ergative, head-final language with pre-nominal relative clauses. A self-paced reading task was used in Experiments 1 and 2, while ERPs were recorded in Experiment 3. We used relative clauses that were ambiguous between an object or subject-gap interpretation and disambiguated later in the sentence. The results of Experiments 1 and 2 showed that SR took longer to read than OR in the critical disambiguating region. In addition, Experiment 3 showed that SR produced larger amplitudes than OR in the P600 window immediately after reading the critical disambiguating word. Our results suggest that SR are not universally easier to process. They cast doubts on universal hypotheses and suggest that processing complexity may depend on language-specific aspects of grammar.

1. Introduction

Cross linguistic investigations are crucial to discover the source of processing asymmetries, and to differentiate between universal processing mechanisms and the impact of grammatical properties of the languages at play on processing. In order to understand the nature of complexity in language processing, a broad sample of different grammars must be studied; the results of this cross linguistic research will reveal the mechanisms at play in language processing at an adequate level of abstraction.

In the domain of sentence comprehension, relative clause processing has been thoroughly investigated in different languages, with various methodologies. A well-established result of these studies is that subject-gap relative clauses (SR), such as (1a) are easier to process than object relative clauses (OR) like (1b).

(1) a. The senator$_1$ [that (e_1) attacked the reporter] admitted the error
b. The senator$_1$ [that the reporter attacked (e_1)] admitted the error.

The sample of languages where this processing asymmetry has been observed consists of a large group of head-initial (SVO), nominative–accusative languages from the Indo-european family, where relative clauses follow the head noun and are typically headed by a complementizer

* Corresponding author. Address: Basque Center on Cognition, Brain and Language, Paseo Mikeletegi 69, 20009 Tenerife, Spain. Fax: +34 943309052.
E-mail address: m.carreiras@bcbl.eu (M. Carreiras).

0010-0277/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2009.11.012

(typically an interrogative element), as the examples in (1). The languages in this group include *English* (e.g., Caplan et al., 2002; Ford, 1983; Gibson, Hickok, & Schutze, 1994; Gordon, Hendrick, & Johnson, 2001; King & Just, 1991; King & Kutas, 1995; Pickering, 1994; Traxler, Morris, & Seely, 2002; Weckerly & Kutas, 1999), *Dutch* (e.g., Frazier, 1987; Mak, Vonk, & Schriefers, 2002, 2006), *French* (e.g., Cohen & Mehler, 1996; Frauenfelder, Segui, & Mehler, 1980; Holmes & O'Regan, 1981), *German* (e.g., Mecklinger, Schriefers, Steinhauer, & Friederici, 1995; Schriefers, Friederici, & Kühn, 1995) and Spanish (Betancort, Carreiras, & Sturt, 2009). This asymmetry has been found with different methodologies such as self-paced reading, eye movements, event related potentials (ERPs) and functional magnetic resonance imaging (fMRI).

Interestingly, there is a second group of languages that have been recently studied, where relative clauses precede their head noun instead of following it, and where there are no interrogative-like words heading the clause. The languages in this group are *Chinese* (e.g., Chien-Jer & Bever, 2006; Hsiao & Gibson, 2003; Lin & Bever, 2006; Lin, submitted for publication), *Japanese* (e.g., Ishizuka, 2005; Ueno & Garnsey, 2008), and *Korean* (e.g., Kwon, Polinsky, & Kluender, 2006; Kwon, Lee, Gordon, Kluender, & Polinsky, submitted for publication). Results from these languages confirm the SR advantage, with the exception of Hsiao and Gibson (2003) as well as Lin and Garnsey (submitted for publication)[1], who report an OR advantage for Chinese, and Ishizuka, Nakatani, and Gibson (2006) who report a similar result in Japanese. However, their experimental materials and results have been recently contested by Chien-Jer and Bever (2006), Lin and Bever (2006) and Kuo and Vasishth (submitted for publication) who also report a SR preference for Chinese and by Kwon et al. (submitted for publication) that did not find an OR advantage in Korean.

Several hypotheses have been offered in the literature in order to explain the processing asymmetry between subject and object relative clauses. Among them we can differentiate a set of hypotheses that predict this effect to be universal, with no appeal to syntactic structure, and those that predict differences across languages depending on the parametric features of the grammar at stake. It is thus important to determine first whether the SR advantage can be seen in languages with properties that diverge from the language-pool that has been previously studied; if the effect is not present in some of these grammars, the first set of hypothesis would be falsified, but if the effect remains across different types of grammars, these set of hypothesis would gain force. One goal of the present study is to explore SR/OR processing asymmetries in Basque, a head-final language with pre-nominal relative clauses, which unlike all previously studied languages is ergative and highly inflected.

Within the group of hypotheses that predict a universal complexity effect without appeal to syntactic structure, we find the Accessibility Hierarchy (e.g., Dowty, 1991; Hale, 2003; Keenan & Comrie, 1977; Keenan & Hawkins, 1987), and the Perspective Shift Hypothesis (e.g., Bever, 1970; MacWhinney, 1977, 1982; MacWhinney & Pleh, 1988), both based on the inherent saliency of subjects relative to objects: the Accessibility Hierarchy claims that grammatical functions are universally ordered in a hierarchy that determines the relative accessibility of a given function; since subjects are placed higher than objects in this hierarchy, subject–object asymmetries should always favour subjects. The perspective shift hypothesis argues that subjects determine the perspective of a clause, and perspective shifts employ processing resources; processing a subject relative clause entails no shift, while object relative clauses induce a shift to a new subject, thus creating a complexity effect.

In the group of structure-dependent hypotheses we find working memory (e.g., Ford, 1983; Frazier & Fodor, 1978; Wanner & Maratsos, 1978), integration cost (e.g., Gibson, 1998, 2000; Hsiao & Gibson, 2003), syntactic strategies such as *Active Filler Strategy* and the *Minimal Chain Principle* (henceforth the AFS and MCP; see Clifton & Frazier, 1989; Frazier & Flores d'Arcais, 1989; Pickering, 1994; Pickering & Barry, 1993, 1991; Stowe, 1989), the simultaneous influence of syntactic and non-syntactic information (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, Tanenhaus, & Kello, 1993), and differences in word-order canonicity (e.g., Bever, 1970; MacDonald & Christiansen, 2002; Mitchell, Cuetos, Corley, & Brysbaert, 1995; Tabor, Juliano, & Tanenhaus, 1997). For an extensive review of these proposals, see Traxler et al. (2002) and Hsiao and Gibson (2003).

On the other hand, there are constraint-based approaches, according to which syntactic and semantic structures are hypothesized to be continuously activated in parallel (cf. Boland, 1997; Gennari & MacDonald, 2008; MacDonald, 1994; MacDonald et al., 1994; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Spivey-Knowlton & Sedivy, 1995; Trueswell, Tanenhaus, & Garnsey, 1994). Comprehension difficulty in this approach emerges from competition between alternative structures partially activated during comprehension. The frequency of the structures in speakers' linguistic experience will determine the activation levels of the different structures, so that infrequent structures in a given configuration would be difficult to activate because of competition with more available frequent structures. Thus, according to this approach the relative difficulty in processing relative clauses may depend on the frequency and the extent of the competition that each structure affords.

A common underlying principle to structure-dependent explanations is the appeal to the relative *distance* between filler-gap dependencies, with increasing distance correlating with increasing complexity. This notion of distance can be characterized in two different ways: in terms of linear distance – the amount of intervening words/terminal nodes, as in the Dependency Locality Theory (DLT, Gibson, 1998, 2000), or in terms of structural distance – the amount of intervening syntactic nodes/projections, as in the Structural Distance Hypothesis (SDH, O'Grady, Miseon, & Miho, 2003). Structural distance between filler and gap is always greater in OR than in SR, because objects are embedded deeper in syntactic structure than subjects. This holds both in head-initial (VO) and in head-final (OV)

[1] They tie the OR preference to particular animacy configurations in the materials.

languages, so that both language groups are predicted to display the same complexity effect by the SDH. Syntactic representations in (2) and (3) schematically illustrate the different predictions made by DLT and SDH for head-initial and head-final grammars. Structures (2a) and (2c) represent a SR in a head-initial and head-final language respectively. Structures (2b) and (2d) illustrate an OR in a head-initial and a head-final language. The subject-gap is higher than the object gap in both types of grammars, regardless of head direction, since the object gap is always embedded deeper in the structure, inside the VP[2]:

(3) a. [$_{RC}$ e_1 Object Verb-rel] filler$_1$ — SR in head final language

b. [$_{RC}$ Subject e_1 Verb-rel] filler$_1$ — OR in head final language

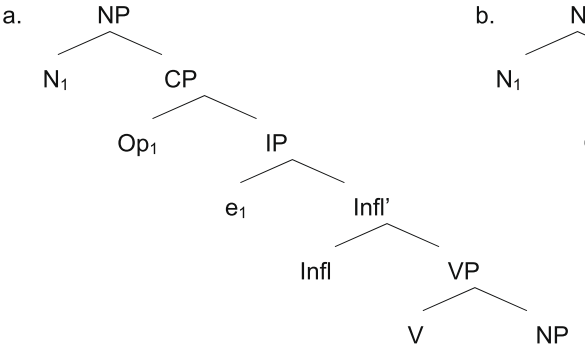
c. filler$_1$ [$_{RC}$ rel e_1 Verb Object] — SR in head initial language

d. filler$_1$ [$_{RC}$ rel Subject Verb e_1] — OR in head initial language

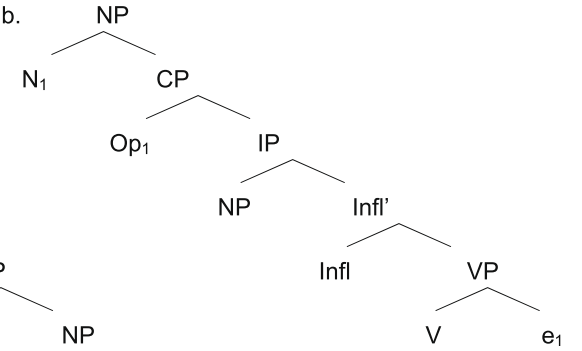
Both notions of distance make similar predictions in head-initial languages, so that it is not possible to distinguish between them by studying only this parametric specification.

(2) Post-nominal SR and OR in a head-initial grammar

a.

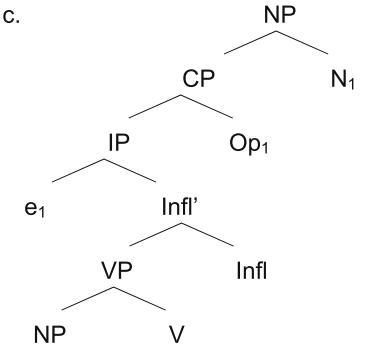


b.

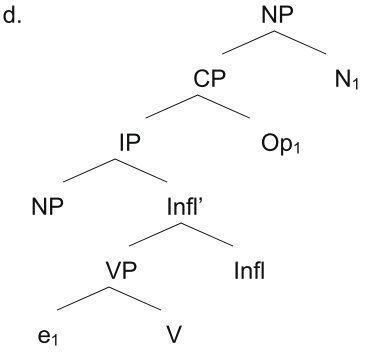


Pre-nominal SR and OR in a head-final grammar

c.



d.



A complexity metric in terms of linear distance, as the DLT, predicts inverse asymmetries depending on the value of the head-parameter: in a head-final language, the linear distance between the gap and the filler of a SR (3a) is greater than the linear distance between the gap and the filler of an OR (3b). However, in a head-initial language, both linear and structural distance between filler and gap is greater in OR (3c) than in SR (3d):

However, in head-final languages each metric yields inverse predictions: the linear gap-filler distance is longer in SR than in OR, but structural distance is greater in OR than in SR. An account based on structural distance (like the SDH) predicts SR to be simpler and easier to process – subjects are higher in the structure of known natural languages (e.g., Baker, 2001; Chomsky, 1957; Clifton & Frazier, 1989; Keenan & Hawkins, 1987; O'Grady, 1997). However, a processing account based on linear distance – intervening material – between filler and gap (like the DLT) predicts SR to be easier than OR only in languages with post-nominal relative clauses, but OR to be easier than SR in languages with pre-nominal relative clauses (e.g., Gibson, 1998, 2000; Hsiao & Gibson, 2003). It is therefore important to clarify what the relative clause processing complexity is in languages with pre-nominal relative clauses to properly adjudicate between these two conceptions of distance.

[2] Recent developments in the *Principles and Parameters* model assume that both subjects and objects establish their thematic relations inside the VP (Koopman & Sportiche, 1991), and then get out of this projection to higher projections (Pollock, 1989). Even under those assumptions, objects always remain hierarchically lower than subjects, which is the critical issue here. For simplicity, we have not included finer syntactic structure in these examples.

A second goal of this paper is therefore to contribute to determine the relevance of linear and structural distance for the processing of long distance dependencies across languages. We investigate the relative complexity of OR versus SR in Basque in three experiments. A self-paced reading task was used in Experiments 1 and 2, while ERPs were recorded in Experiment 3. Relative clauses in Basque, as in Chinese, Korean and Japanese, are temporally ambiguous, and speakers might not know a RC is coming until they find the inflected verb of the RC, the last word before the head noun. There are no interrogative pronouns, only a complementizer morpheme *-(e)n* attached as a suffix to the inflected verb, like the form *ditu-en* shown in (4), where *ditu* is an auxiliary verb, and *-en* is the complementizer that indicates the presence of a RC. Experimental sentences were constructed using the *-ak* ending, which yields an ambiguity between singular transitive subject and plural object, so that the relative clauses were ambiguous between a SR or OR reading until the last word of the main sentence ((4a) and (4b), respectively)..

(4) a. [$_{RC}$ e$_1$ [$_{VP}$ irakasle-ak aipatu] ditu-en] ikasle-a-k$_1$ lagun-ak ditu
[$_{RC}$ e$_1$ [$_{VP}$ teacher-pl mentioned] has-rel] student-sg-S$_1$ friend-pl has
"The student [that e mentioned the teachers] has friends"
b. [$_{RC}$ irakasle-a-k [$_{VP}$ e$_1$ aipatu] ditu-en] ikasle-ak$_1$ lagun-ak dira
[$_{RC}$ teacher-sg-S [$_{VP}$ e$_1$ mentioned] has-rel] student-pl$_1$ friend-pl are
"The students [that the teacher mentioned e] are friends"

The *-ak* ending illustrated in the examples in (4) is ambiguous between two morphological classes: (i) a plural absolutive Noun Phrase (i.e. a plural object or a plural intransitive/thematic subject), or (ii) a singular ergative Noun Phrase (i.e. a transitive/agentive subject). In each of these two cases, the *-ak* sequence has a different morphological structure. When interpreted as belonging to (ii), the class of singular ergative NPs, it consists of a determiner *-a* "the" and an ergative case marker *-k* attached, as illustrated in (4b) for instance, where the singular transitive subject *irakasleak* "the teacher" consists of the noun *irakasle* "teacher", plus the singular determiner *-a* and the ergative case marker *-k* (*irakasle-a-k* "the teacher"). When interpreted as belonging to (i), the class of absolutive plural NPs, there is only one morpheme, the plural determiner *-ak* "the$_{pl}$", and no case marker attached, for absolutive is morphologically unmarked. This can also be seen in (4b), where the plural intransitive (hence absolutive) subject *ikasleak* consists of the noun *ikasle* "learner, student" plus the plural determiner *-ak* (*ikasle-ak*), and no overt case marker. Similarly, in (4a), *irakasle-ak* is a plural object (hence also absolutive) "the teachers", whereas *ikasle-a-k* is a transitive subject "the student" (see Laka, 1996, for details on case morphology).

Match or mismatch between the case-marking on the head noun of the RC and the gap inside the RC has been argued to have an effect in processing difficulty between the two types of RCs (Sauerland & Gibson, 1998). There is no potential for a case-mismatch confound in our experimental sentences, because they always involve a case-match configuration: on the one hand, OR sentences have absolutive case-marked subjects as head nouns, so both the gap and the head noun carry absolutive case, as in (4b), where the object gap corresponds to an absolutive case-marked NP, and the head noun also carries absolutive case; on the other hand, SR sentences have ergative-marked subjects as head nouns, so both the gap and the head have ergative case, as in (4a), where the subject-gap corresponds to a NP that would carry ergative case, the same case as its head noun in the main clause. That is, in both types of experimental sentences the similarity of the case borne by the gap and head NP is absolute.

There are no differences between the two RCs in terms of storage resources (Gibson, 2000), because the number of unresolved dependencies is the same in the two clauses. Both in the OR (4b) and in the SR (4a), only one head needs to be postulated to generate a grammatical sentence: an inflected verb that would yield a well-formed intransitive main clause, because given the ambiguity of the first overt phrase, a grammatical intransitive sentence can be completed by simply adding a verb: *ikasle-ak datoz* "the students arrive". Secondly, we can also discard the effect of word-order canonicity (MacDonald & Christiansen, 2002) because the relative clauses do not follow the SOV canonical order of the language; (4b) presents a SVO order, and (4a) an OVS order, so that a processing asymmetry cannot be explained by the canonicity hypothesis. The structural distance between the subject-gap and the filler is shorter than the structural distance between the object gap and the filler, because the later is lower in syntactic structure (cf. structures (2c) and (2d)). Thirdly, linear distance or integration cost is higher for SR than for OR in these sentences: in the OR (4b), the number of intervening words is two (verb and inflected auxiliary), whereas in the SR (4a) the number of intervening words is three (object, verb and inflected auxiliary). Finally, similarity or differences in frequency of use for the two relative clauses will be measured in a corpus analysis to investigate whether frequency of use can account for processing differences.

In sum, if an object relative clause preference is obtained it will suggest that (a) SRs are not universally, inherently easier to process than OR, so that accounts based on the inherent or universal properties of subjects would be ruled out, and that (b) integration cost in terms of structural distance does not explain processing preferences across grammars. Moreover, if an OR advantage were to obtain, given the SR advantage results in Japanese and Korean, the explanatory role of linear distance between filler and gap would require further scrutiny.

2. Experiment 1

Experiment 1 used a self-paced reading task – moving window – to examine reading times for the experimental stimuli, in order to determine whether they show that SR are harder to process and therefore take longer to read than OR or vice versa. Self-paced reading tasks such as

the moving window, which we will describe in the method, have been widely used in the sentence comprehension field to investigate the processing load associated syntactic parsing (see Mitchell, 2004 for a review of the assumptions underlying this technique).

2.1. Method

2.1.1. Participants

Fifty-four native Basque speakers, all undergraduate students from the University of the Basque Country (UPV/EHU), took part in this data collection, which was carried out in ELEBILAB, the Psycholinguistics Laboratory in Vitoria-Gasteiz. They received 3 € in exchange for their participation.

2.1.2. Materials

Fifty pairs of sentences were created for this experiment. Each sentence consisted of six words. The last word in each sentence (the inflected verb) was the one that indicated the structure was an embedded clause, and disambiguated the RC for a SR or an OR. Therefore, the two sentences in a pair only differed in the inflected verb of the main sentence (see examples 5a and 5b). In addition, in order to control for the possible influence of differences in the plausibility of the RCs in the two critical conditions (OR versus SR), we conducted a questionnaire to assess the naturalness of each of the RCs. The items consisted of the simple clauses that formed each RC with a SOV presentation structure (subjects and objects were counterbalanced in the two lists). Twenty native speakers, all of them undergraduate students at the University of the Basque Country (UPV/EHU) that did not participate in any of the experiments rated each clause in a 1-to-7 scale, one being totally unnatural and seven totally natural. The naturalness ratings in the two lists were highly similar (5.8 and 5.6), revealing that there were no naturalness differences between the sentences ($p > .27$ in a t-test).

5a. (Subject relative clause).
Irakasleak aipatu dituen ikasleak lagunak ditu.
[e_1 irakasle-ak aipatu ditu-en] ikasle-a-k_1 lagun-ak ditu.
[e_1 teacher-pl mentioned has-rel] student-sg-S_1 friend-pl has.
"The student that mentioned the teachers has friends"

5b. (Object relative clause).
Irakasleak aipatu dituen ikasleak lagunak dira.
[irakasle-a-k e_1 aipatu ditu-en] ikasle-ak_1 lagun-ak dira.
[teacher-sg-S e_1 mentioned has-rel] student-pl_1 friend-pl are.
"The students that the teacher mentioned are friends."

These sentences were divided in two lists following a counterbalanced design. Each list contained 25 SR and 25 OR sentences. We also created a set of 115 unambiguous filler sentences, so that the percentage of SR or OR was 15% in each list. Different participants were randomly assigned to each list. Trial order repetition effects were avoided randomizing the presentation order for each participant.

2.1.3. Procedure

Participants were individually presented one list of materials in a silent and correctly illuminated room. The stimuli presentation and data collection was done with the Linger software by Doug Rohde (see http://tedlab.mit.edu:16080/~dr/Linger/) installed in PC computers with CRT monitors. Participants were initially presented with a string of dashes. Each time the F button of the keyboard was pressed one word was unmasked, keeping the rest of the sentence masked. Only one word was unmasked each time. In 50% of the sentences comprehension questions were displayed, and participants had to press one of two buttons in order to make a choice. The target region in the present experiment was always the last reading region (the sixth word in the experimental sentences), since this was the disambiguation region. The whole session, including a practice of eight sentences, lasted less than 25 min.

2.2. Results and discussion

Reading times in the critical disambiguation window beyond or above the 2.5 standard deviation cutoff values were not included in the analyses (less than 1.8% of the data). Participant and item one-factor ANOVAs with two levels (Type of RC: SR, OR) were carried out. Mean reading times are displayed in Fig. 1. Critical regions that included the word that disambiguated the RC in favour of an OR were read much faster than those regions with the word that disambiguated for a SR (a 115 ms difference), $F1(1, 52) = 8.24$, $p < .01$; $F2(1, 48) = 19.20$, $p < .01$. No differences were observed in none of the previous regions (all $Fs < 1$).

The results were clear cut. Object relatives were easier to process than subject relatives. Nonetheless, it could be argued that two other hypotheses could account for the present results. (1) Longer reading times for subject relative clauses might reflect a garden path effect because readers end up reading a subject relative clause but they were expecting an object relative clause. (2) Long reading times in the subject relative disambiguation might reflect a garden path effect for the *-ak* singular/plural morphological ambiguity that is also resolved when reader reach the main verb. The co-existence of the morphological and the syntactic ambiguity would prevent to exclusively attribute the effects to subject versus object relative clause processing difficulties.

The first alternative hypothesis is a consequence of the fact that subject relatives are harder to process than object relatives with the current set of materials, because the same words are included for both object and subject relative clauses up to the disambiguation point. Therefore, the same plausibility or interpretation bias has to be assumed up to the disambiguation point for the two types of sentences. Thus, assuming that sentence processing is incremental, if there is a preference for object relative clauses in Basque, we should expect that readers get into a garden

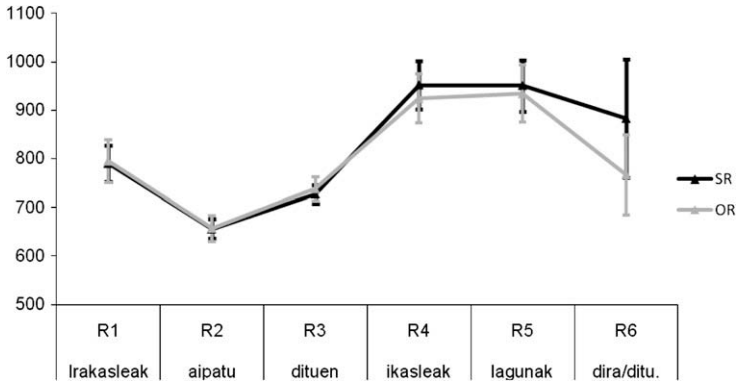


Fig. 1. Reading times for the different regions of the sentences containing the subject and object relative clauses in Experiment 1.

path when they read a subject relative clause because they expected and object relative clause. In addition, it is parsimonious to assume that the preference for object relative clauses is all along the way from the beginning of the sentences up to the disambiguation point, given that the same words are present in the two sentences.

The second alternative hypothesis is not a problem either, because the resolution of two ambiguities (morphological and syntactic) lead to the opposite pattern of results. Disambiguation to plural takes more time, but disambiguation to plural is only consistent with a subject relative clause interpretation. Thus, while the morphological disambiguation predicts a subject preference, the syntactic pattern predicts an object preference. More clearly, in Basque it is the subject relative that contains the singular head noun interpretation (see 4a and 5a) and the object relative that contains the plural head noun interpretation (see 4b and 5b).

Therefore, because the processing of the plural form takes longer, the expectation is to find a subject relative advantage if the effect is due to the morphological processing of the plural form because the subject relative disambiguation corresponds with a singular disambiguation. Instead, and despite the fact that the object relative contains a plural noun head interpretation, results show faster reading times for object relative disambiguation.

In sum, the empirical evidence from Experiment 1 shows that subject relative clauses are harder to process, which is against the predictions of a universal account of structural distance computation. This hypothesis predicts that object relative clauses are less accessible and harder to extract than subject relative clauses. However, as we will argue in the General Discussion, this result can be accounted by the integration cost hypothesis (see Hsiao & Gibson, 2003). This is an important finding that needs to be replicated because its very relevant theoretical implications. In addition, it should be noted that the critical disambiguating region was composed of the last word in each of the sentences, and this could have been influenced by wrap-up effects associated with sentence final integration processes (e.g., Just & Carpenter, 1980). In Experiment 2 we included a larger set of sentences and added an extra word at the end of each of the sentences (after the critical disambiguating verb), so that the disambiguation for object or subject relative clauses occurred always in the penultimate word.

3. Experiment 2

3.1. Method

3.1.1. Participants

A different group of 22 undergraduates from the University of the Basque Country took part in this experiment. All of them had normal or corrected-to-normal vision and were native Basque speakers.

3.1.2. Materials

Eighty experimental sentence pairs containing a RC were constructed for this experiment. Each sentence was seven-word-long. The two sentences in each pair only differed in the sixth word (the inflected verb of the main sentence), that disambiguated the RC for a SR or for an OR (see examples 6a and 6b). Thus, the disambiguating word was never the last word of the sentence. These 80 sentence pairs were split in two lists, so that there were 40 SR and 40 OR in each list. Among these 80 sentence pairs, all the 50 sentence pairs from Experiment 1 were included. A set of 230 filler sentences of the same length was also included in each list, so that the percentage of experimental sentences was 25%. Participants were randomly assigned to one of the lists. The trials were presented in a different random order to each participant.

In order to further control for the possible influence of differences in the plausibility of the RCs in the two critical conditions (OR versus SR), a second plausibility normative survey was also carried out for the entire sentences (not only for the embedded RCs). To this end, a group of 38 different undergraduate students from the University of the Basque Country completed a plausibility questionnaire in which they were asked to rate each sentence from Experiments 2 and 3 in a 1-to-7 scale for their naturalness (1 = not plausible, 7 = totally plausible). (Note that the sentences from Experiment 2 also include those used in Experiment 1). The obtained scores were very similar both for the sentences that disambiguated for subject RCs and for

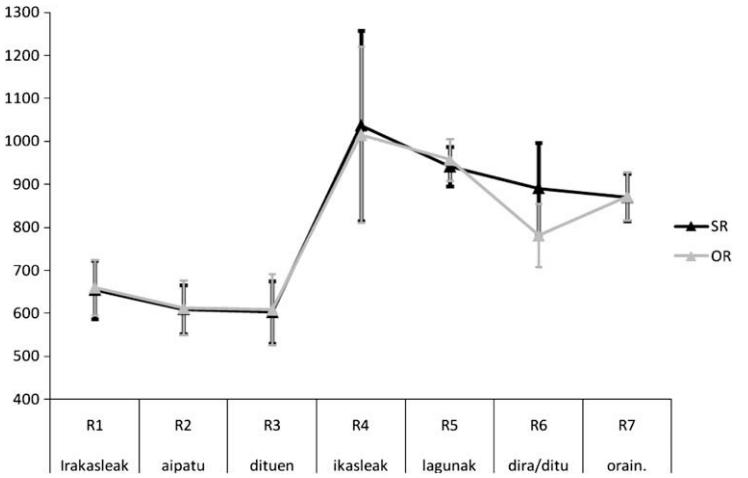


Fig. 2. Reading times for the different regions of the sentences containing the subject and object relative clauses in Experiment 2.

those that disambiguated for object RCs (mean punctuations of 4.7 and 4.7 in both conditions; $p > .89$ in a *t*-test).

6a (Subject relative clause)
Irakasleak aipatu dituen ikasleak lagunak ditu orain.
[e_1 irakasle-ak aipatu ditu-en] ikasle-a-k_1 lagunak ditu orain.
[e_1 teacher-pl mentioned has-rel] student-sg-S_1 friend-pl has now.
"The student that mentioned the teachers has friends now".

6b. (Object relative clause)
Irakasleak aipatu dituen ikasleak lagunak dira orain.
[irakasle-a-k e_1 aipatu ditu-en] ikasle-ak_1 lagun-ak dira orain.
[teacher-sg-S e_1 mentioned has-rel] student-pl_1 friend-pl are now.
"The students that the teacher mentioned are friends now".
"The students that the teacher mentioned are friends now".

3.1.3. Procedure

The same procedure as for Experiment 1 was followed. The target region in the present experiment was the sixth word in the experimental sentences. The whole session lasted less than 45 min.

3.2. Results and discussion

Reading times in the critical disambiguation window beyond or above the 2.5 standard deviation cutoff values were not included in the analyses (less than 4% of the data). Participant and item one-factor ANOVAs with two levels (Type of RC: SR, OR) were carried out. Mean reading times are displayed in Fig. 2. Critical regions that included the word that disambiguated the RC in favour of an OR were read much faster than those regions with the word that disambiguated for a SR (a 109 ms difference), $F1(1, 20) = 4.81$, $p < .05$; $F2(1, 78) = 10.79$, $p < .01$. No differences were observed in none of the previous regions (all Fs < 1).

The results of Experiment 2 replicated the findings obtained in Experiment 1: object relative clauses were easier to process than subject relative clauses. Hence, Experiment 2 constitutes a replication of the previous findings with a larger set of sentences. Furthermore, in contrast to the sentences in Experiment 1, this set of sentences included an extra word after the region of disambiguation, so that the obtained results cannot be attributed to sentence final wrap-up effects (e.g., Just & Carpenter, 1980).

4. Corpus study

The differential comprehension difficulty for subject and object relative clauses observed in Experiments 1 and 2 could emerge from the activation of several competing structures ultimately derived from distributional patterns of language. Probabilistic approaches would suggest that differences in plausibility, animacy and/or frequency between subject and object relative clauses would be at play. However, we already discarded plausibility as a factor in the two previous experiments. On the other hand, although several studies have pointed out that comprehension difficulty in object relatives in languages such as English and Dutch varies with the animacy configuration of the nouns involved (e.g., Mak et al., 2002; Traxler et al., 2002), in the present case both antecedents were animate, and we are finding that subject relatives are harder. A third possibility could be the differential frequency of usages of subject and object relative clauses. If this is the cause, it would even account not only for the pattern of results in Basque, but also for the fact that this is a reversed pattern of that obtained in English. In other words, if object relative clauses are more frequency than subject relative clauses in Basque the differential activation depending on frequency of usage will be a straightforward explanation for the results found in Experiments 1 and 2 within the constrain satisfaction framework. To investigate this account we carried out a corpus analysis in Basque.

A corpus study was carried out to check for frequency differences of each of the structures (OR versus SR). To this

end, we selected a Basque corpus consisting of 55,000 words, and computed the relative frequency of subject and object RCs. Clearly, if the number of OR is found to be significantly greater than the number of SR, our results could be explained simply in terms of frequency or familiarity of the structures.

The corpus employed for this frequency count is EPEC (Aduriz et al., 2006). EPEC is a 55,000-word sample collection of written standard Basque. Half of it has been extracted from the Statistical Corpus of 20th Century Basque (http://www.euskaracorpusa.net). The other half was extracted from *Euskaldunon Egunkaria* (http://www.egunero.info), a daily newspaper written entirely in Basque. The Statistical Corpus of 20th Century Basque is a reference corpus of Basque including 4658,036 word-forms. It was created by UZEI (http://www.uzei.com), a non-profit organization devoted to making Basque language suitable for any specialized field. A sub-corpus of about 25,000 word-forms was extracted from this large corpus in order to build EPEC. Texts written in standard Basque, corresponding to the last period (1991–1999) and belonging to both literary and non-literary prose, were chosen for this purpose. The second part of EPEC consists of several articles extracted from the *Euskaldunon Egunkaria* written in the second half of 1999 and in 2000. The articles were chosen so that they covered an assorted range of topics (economics, culture, entertainment, international, local, opinion, politics, sports...). This corpus was created and morphosyntactically tagged by the IXA group, the computational linguistics research team from, Department of Computer Languages and Systems at the University of the Basque Country (Donostia-San Sebastián), who provided us with a list of all relative clauses in the corpus.

In the cited corpus, a total of 625 pre-nominal RCs were found. From the total, 399 were SR (approximately 64%), while only 226 were OR (approximately 36%). Thus, the frequency explanation does not seem suitable for accounting for the present findings (note that, if any, the influence of the frequency of appearance should have led to faster or less costlier recognition of SR as compared to OR). The results are not consistent with the hypothesis that frequency plays a role in determining the alternative available structures since these distributional patterns are no consistent with the reading patterns. The main prediction of the constrain satisfaction approach is that there is a direct link between frequency of structures and on-line comprehension difficulty, what does not seem to happen in the current situation.

Some other researchers (e.g., Gibson, 1998; Gordon, Hendrick, & Johnson, 2004; Grodner & Gibson, 2005) have endorsed a similar view arguing that frequency information, a major constraint in ambiguity resolution, cannot account for processing difficulty in object/subject relative clauses.

5. Experiment 3

Experiment 3 investigated ERPs in response to SR and OR in Basque. ERPs are averages of brain electrical activity time-locked to some external or internal event and classified according to their polarity (i.e., positive or negative deflections in the waveform), the time of their onset or peak occurrence in milliseconds, and their topographical distribution across the scalp. ERP studies have provided crucial information, with an exquisite time resolution, about language processing (for review see Kutas, Van Petten, & Kluender, 2006; Osterhout, McLaughlin, & Bersick, 1997). Syntactic processing has been associated with two separate ERP components. A *left anterior negativity* (LAN) has been elicited by syntactic (e.g., phrase structure, subcategorization) and morphosyntactic (e.g., subject–verb, article–noun and antecedent–pronoun agreement) violations. This component has an anterior distribution, starting as early as 250 ms, and larger amplitude in the left hemisphere (e.g., Barber & Carreiras, 2005; Coulson, King, & Kutas, 1998; Friederici, Pfeifer, & Hahne, 1993; Kluender & Kutas, 1993). The second component associated with syntactic anomalies and non-preferred syntactic structures is a large positive wave that onsets at about 500 ms after presentation of the anomalous word and persists up to approximately 900–1000 ms. (the *P600 effect*/Syntactic Positive Shift; see Barber & Carreiras, 2005; Coulson et al., 1998; Hagoort, Brown, & Groothusen, 1993; Osterhout & Holcomb, 1992; Osterhout & Mobley, 1995). In particular, the P600 or Syntactic Positive Shift (SPS) has been related to the processes of revision and repair in sentence processing. In sum, the LAN is elicited only after certain types of violations, while the P600 is elicited after most syntactic violations and even by syntactically ambiguous structures without ungrammaticality.

Some previous studies have investigated relative clause processing in English using ERPs. For instance, King and Kutas (1995) investigated the processing of the subject and object relative clauses and found a greater LAN to main clause verbs of object as compared with subject relative structures; that is in the verb after the relative clause. In addition, they compared the two sentences not only a word-by-word basis but across their entire extent. The reason behind is that subject and object relative clause were well controlled both for length and identity of the lexical items involved, although they were comparing different words at each particular point inside the relative clause. This latter analysis revealed a frontal negativity for object relative clauses as compared to subject relative clauses that became noticeable at the start of the relative clause, it was lost in the later part of the relative clause, and appeared again during the processing of the main clause. This slow brain potential effect was interpreted as reflecting greater demands on working memory. In addition, Weckerly and Kutas (1999) contrasted two object relative types that were syntactically and lexically identical and varied only in the order of the component animate and inanimate nouns [Inanimate/Animate) versus Animate (Inanimate)]. The ERPs to the main clause verbs in A(I) sentences showed both a LAN effect and an P600 relative to those in I(A) sentences. In other words, between 200 and 500 ms, the response to main clause verbs in A(I) sentences was more negative than to those I(A) sentences.

In sum, previous knowledge about the eliciting conditions of ERP effects after syntactic violations, and in particular previous evidence with subject and object relative clauses, together with the evidence presented in Experi-

ments 1 and 2, will allow us to predict that ERPs after the critical disambiguating word should elicit at least a difference in the P600 window or differences in the LAN/P600 windows. In particular, given the results obtained in Experiments 1 and 2 we expect to find a P600 or a LAN/P600 effect showing larger amplitude for Subject Relative clauses, which were the most difficult to process.

5.1. Method

5.1.1. Participants

Twenty-two different native Basque speakers, all undergraduate students at the University of the Basque Country, took part in this data collection, which took place in ELEBI-LAB, the Psycholinguistics Laboratory in Vitoria-Gasteiz. They received 15 € in exchange for their participation.

5.1.2. Materials

The same materials (experimental and filler sentences) as in Experiment 2 were used.

5.1.3. Procedure

A fixation point ("+") appeared at the center of the screen and remained there for 1000 ms. This fixation point was followed by a blank screen interval of 300 ms, then the sentence was displayed word by word. Each word appeared for 300 ms and was followed by a 300-ms blank interval. Participants were instructed to read the sentences for comprehension. The inter-trial interval varied randomly between 1500 and 2000 ms. Two counterbalanced lists were finally used for the experiment. Before starting the experimental phase 8 warm-up practice trials were presented to the participants.

5.1.4. EEG recording and analyses

Scalp voltages were collected from 58 Ag/AgCl electrodes which were mounted in an elastic cap (ElectroCap International, Eaton, USA, 10-10 system). The right mastoid earlobe was used as reference (see Fig. 3). Eye movements and blinks were monitored with two further electrodes providing bipolar recordings of the horizontal and vertical electrooculogam (EOG). Inter-electrode impedances were kept below 5 KΩ. EEG was filtered with an analogue band-pass filter of 0.01–50 Hz and a digital 30 Hz low-pass filter was applied before analysis. The signals were sampled continuously throughout the experiment with a sampling rate of 500 Hz, and digitally re-referenced to linked mastoids.

Epochs of the EEG corresponding to 800 ms after target word onset presentation were averaged and analyzed. Baseline correction was performed using the average EEG activity in the 200 ms preceding the onset of the target stimuli as a reference signal value. Following baseline correction, epochs with simultaneous artefacts in at least 10 channels were rejected. This resulted in the exclusion of approximately 16% of the trials. Separate ERPs were formed for each of the experimental conditions, each of the subjects and each of the electrode sites.

Six regions of interest were computed out of the 61 electrodes, each containing the mean of a group of electrodes. The regions were (see electrode numbers in Fig. 1): left-anterior (F1, F3, F5, C1A, C3A, C5A), left-central (C1, C3, C5, C1P, C3P, TCP1), left-posterior (P1, P3, P5, P1P, P3P, CB1), right-anterior (F2, F4, F6, C2A, C2A, C2A), right-central (C2, C4, C6, C2P, C4P, TCP2), right-posterior (P2, P4, P6, P2P, P4P, CB2).

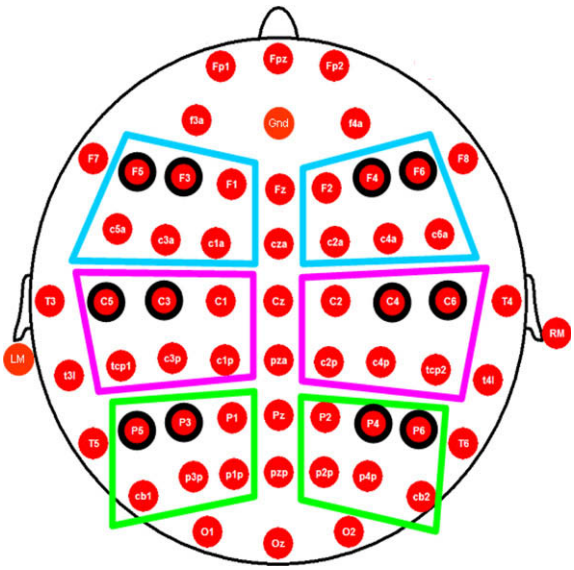


Fig. 3. Schematic flat representation of the 59 electrode positions from which EEG activity was recorded (front of head is at top). The electrodes contributing to the analyses are those grouped in the six critical regions. Electrodes rounded in black are those displayed in Fig. 4.

Mean amplitudes were obtained for a time window between 450 and 700 ms post-stimuli. For each window, a repeated-measures ANOVA was performed, including *electrode regions* (anterior, central and posterior), *hemisphere* (left/right) and the *type of sentence* (subject relative clause versus object relative clause) as factors. Where appropriate, critical values were adjusted using the Greenhouse–Geisser (1959) correction for violation of the assumption of sphericity.

5.2. Results and discussion

The ERP grand averages, time-locked to the onset of the critical disambiguating word for subject and object relative clauses, are represented in Fig. 4. Visual inspection reveals differences in the electrophysiological responses between the two conditions only in the window between 450 and 700 ms that corresponds with the window in which the P600 it is usually found. Subject relative clauses showed larger positivity amplitude in the P600 window than object relative clauses. The topographical distribution of the effect is represented in Fig. 5. No differences appeared in the 300–450 ms window in which the LAN has been usually reported.

The ANOVA with the average values of the 450–700 ms time epoch showed a marginal effect [$F(1, 21) = 3.5$, $p < .07$] and a significant interaction of the factors type of relative clause and hemisphere [$F(1, 21) = 6.07$, $p < .05$]. Significant differences between subject and object relative clauses were obtained in the left hemisphere [$F(1, 21) = 4.7$,

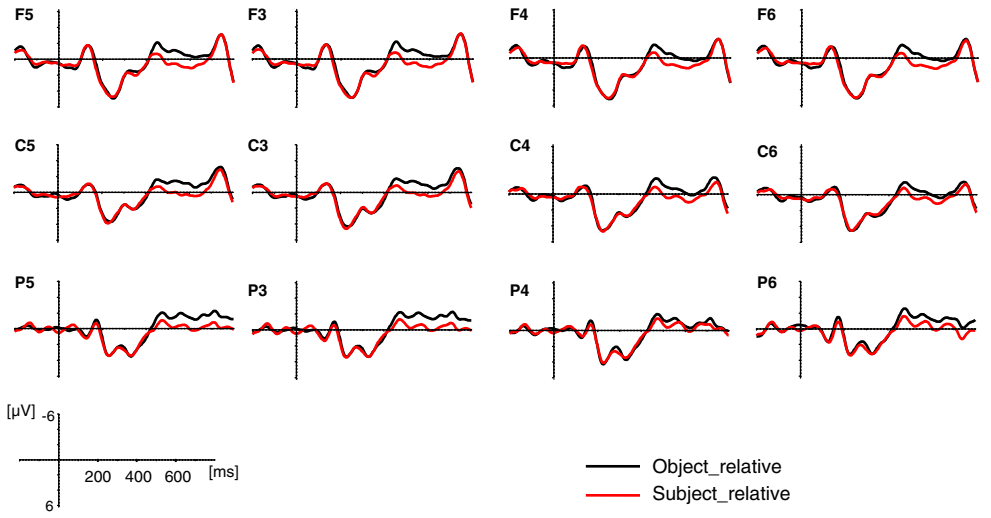


Fig. 4. Grand average ERPs corresponding to the object and subject relative clauses in representative electrodes of the six regions of interest.

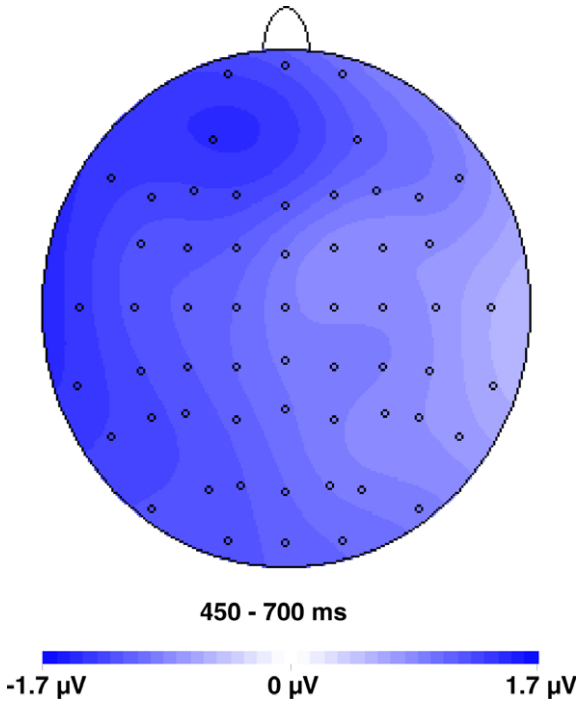


Fig. 5. Topographical distribution of the object minus subject relative clause effect in the 450–700 ms window.

$p < .05$], but not in the right hemisphere [$F(1, 21) = 2.3$, $p = .13$].

The results show that in the P600 window the amplitude of the SR is more positive going that of the OR sentences. Therefore, once more, the present results seem to show that in Basque, SR are harder to process than OR. While the time-course of the effect is within the range reported for the P600 (see Barber & Carreiras, 2005; Carreiras, Salillas, & Barber, 2004; Coulson et al., 1998; Hagoort et al., 1993; Osterhout & Holcomb, 1992; Osterhout & Mobley, 1995) the topography is not according to the standard description. The P600 is mostly described with a posterior distribution. However, in the present case, the distribution of the P600 is widely spread in the left hemisphere. However, this is not the only time this distribution has been found (see, Carreiras et al., 2004; Frederici, Hahne & Saddy, 2002; Hagoort, Brown, & Osterhout, 1999; Kaan & Swaab, 2003; Osterhout & Holcomb, 1992; Van Berkum, Hagoort, & Brown, 1999). Given the atypical scalp distribution, and even the morphology of the wave, it could be argued that instead of a P600 in which SR are showing a larger positivity, the data could be interpreted as an N400 or even a widely distributed LAN in which OR show a larger negativity. Several reasons make us think that this is not the case. Firstly, the time windows reported for the N400 and the LAN do not usually last as in the present case aside from the fact that and N400 interpretation would imply to postulate some semantic anomaly detection. Secondly, the N400 and the LAN usually start earlier. Thirdly, the previous data obtained in Experiments 1 and 2 conflict with this interpretation, because the results of Experiments 1 and 2 show that OR are easier to process than SR. It is also important to note that while the P600 effect with a posterior distribution was generally found for ungrammatical sentence continuations (Coulson et al., 1998), a more frontal/broad distribution of the positivity has been reported for non-preferred continuations (e.g., Frederici, Hahne & Saddy, 2002; Hagoort et al., 1999; Osterhout & Holcomb, 1992; Van Berkum et al., 1999). In fact, Kaan and Swaab (2003) found a frontally distributed P600 effect when complex ambiguous sentence structures were compared with simple grammatically correct unambiguous sentence structures. In addition, Carreiras et al. (2004) found a left distribution for non-preferred continuations.

Finally, it could also be argued that the present ERP data could be interpreted as being consistent with the English results reported by King and Kutas (1995; see also Weckerly & Kutas, 1999). King and Kutas (1995) showed a frontal positivity larger for object relative clauses in a multiword analysis starting at the beginning of the relative clause,

and a LAN when reading the main verb after finishing the relative clause. A LAN effect was also found by Weckerly and Kutas (1999) at the main verb after the relative clause. However, there are various reasons to believe this is not the case. Firstly, the frontal negativity effect reported by King and Kutas (1995) in the multiword analyses comes from the comparison of the early region of the two relative clauses that contain different words (e.g., "the senator" versus "harshly attacked") in the following examples of object and subject relative sentences (*the reporter who the senator harshly attacked admitted the error* versus *the reporter who harshly attacked the senator admitted the error*). Thus, while these findings are consistent with the notion of anterior negativity as a reflection of processing complexity and working memory operations, these effects may also reflect ERP differences to items of different lexical categories. Keep in mind that we are reporting differences during reading of the critical disambiguating word, which is matched for length and frequency in the two conditions. Secondly, when they carry out the word-by-word analyses on the mean amplitude between 300 and 500 ms post-onset for the specific lexical items to examine some of the processing differences noted at the multiword level in a shorter time scale, they do no find differences in the relative clause. They only found a LAN negativity effect very late in the sentence, during reading the main verb. Thirdly, as similar LAN effect was observed by Weckerly and Kutas (1999) at the main verb but they manipulated the animacy of the nouns that served as the relative and main clause subjects only in object relative sentences in a no completely balanced design. Finally, the time windows reported in King and Kutas (1995) for the frontal negativity and that of the present experiment do not overlap.

In sum, because the larger frontal negativity effects found in English for object relative clauses cannot be unequivocally attributed to the manipulation of object versus subject relative clauses, and because the windows reported in English and in Basque do not overlap, we think the two sets of results are hard to compare. Therefore, it would be very difficult to sustain that these data are comparable to the data obtained in English.

6. General discussion

The results of Experiments 1 and 2 showed that SR took longer to read than OR in the critical disambiguating region. In addition, Experiment 3 showed that SR produced larger amplitudes that OR in the P600 window immediately after reading the critical disambiguating word. Thus the three experiments reported here suggest that SR in Basque are harder to process than OR. These results are incompatible with accounts of the relative ease of SR/OR processing that rely on the inherent saliency of subjects, such as the Accessibility Hierarchy (e.g., Keenan & Comrie, 1977), or the Perspective shift hypothesis (e.g., Bever, 1970; MacWhinney, 1977, 1982; MacWhinney & Pleh, 1988), because these accounts predict that SR are easier to process than OR regardless of the specifics of linguistic structure.

Neither are these results compatible with a constrain satisfaction account, since plausibility differences or frequency differences can account for the results. In addition, both possible antecedents were animate, so animacy configuration does not seem to be at play to explain the larger comprehension difficulty for subject relative clauses.

Turning to hypotheses that appeal to specific aspects of the structures to be processed, a common underlying factor to these accounts is that greater distance between filler and gap increases processing complexity. Distance can be characterized in terms of linearity, that is, intervening material (Gibson, 1989), or in terms of the amount of projections in syntactic structure as in the Structural Distance Hypothesis (SDH) (O'Grady et al., 2003). Our results are incompatible with structural distance, which predicts a SR advantage, but compatible with linear distance because the subject-gap in (2a) is linearly further from its filler than the object gap in (2b). However, results form other studies on languages with pre-nominal relative clauses yield a different picture.

Hsiao and Gibson (2003) also reported OR advantage in Chinese and argued for a linear distance account of filler-gap dependencies, but their experimental materials and results have been contested by Chien-Jer and Bever (2006) and Lin and Bever (2006), who report a SR preference for this language. For Japanese, and Korean, Ishizuka (2005), Ueno and Garnsey (2008) and Kwon et al. (2006, submitted for publication) respectively report that SR are processed faster than OR, thus favouring either the structural distance over the linear distance account in the case of Ishizuka (2005) and Ueno and Garnsey (2008), or favouring the Accessibility Hierarchy in the case of Kwon et al. (2006, submitted for publication). If results from languages with pre-nominal relative clauses do not converge, some other factor must be at play. In Ishizuka (2005), the SR sentences were unambiguous, but the OR materials were temporally ambiguous between a main clause with an unexpressed subject and an object relative clause. This temporal ambiguity could presumably trigger a reanalysis for the OR materials that artificially increased its processing complexity, as the author discusses. The materials in our study do not present this potential problem: both SR and OR sentences were equally ambiguous until the same word of the main sentence (i.e., the last word in Experiment 1, and the penultimate word in Experiments 2 and 3). On the other hand, Kwon et al. (2006,submitted for publication) concluded that even in head-final languages, subject-gaps enjoy processing advantage, so that the effect is claimed to be universal. This conclusion is however not granted by our results. How can these apparently conflicting results be reconciled?

If neither linear nor structural distance can provide a coherent account of processing asymmetries across languages, it is necessary to consider other language specific factors that could be influencing the results obtained. There are two possible ways of reconciling SR advantage results from Chinese, Japanese and Korean, and OR advantage results from Basque, both of which involve language-specific properties. One of them is the nature of the gap in the relative clause: as discussed by Ishizuka (2005), Ueno and Garnsey (2008) and Kwon et al. (2006,submitted for publication), many linguistic analysis of Chinese, Japanese and Korean RCs argue that the gap in the relative clause is

a null pronoun (*pro*), which does not result from syntactic movement. This explains a number of idiosyncratic properties displayed by relative clauses in these languages: insensitivity to island effects, gapless relative clauses, and lack of weak cross over effects. In fact, while Ishizuka (2005) assumes that the gaps in the relative clauses are null pronominals, Kwon et al. (2006) assume they are traces of movement. Gaps in Basque relative clauses have been consistently argued by linguists to be *traces* of movement of the same type found in English, French or Spanish (Artiagoitia, 1990,1992; Oyharçabal, 1987). Whether the relation established between a null pronominal and a coreferent element involves the same processing mechanisms as the linking of a trace of movement with an antecedent is an issue that requires further research (see Kown et al., 2006 for a discussion); it is well known, for instance, that pragmatic factors enter in the computation of pronominal reference, while trace-antecedent relations are fully determined by syntactic locality (Chomsky, 1986; Lasnik & Uriagereka, 1988). If the type of gap has an effect on the processing strategy employed, this could account for the conflicting results in the sample of languages with pre-nominal RCs, provided that linear distance is crucial for trace-antecedent integration in the case of languages with relative clauses containing traces of syntactic movement (English, Basque), but not for the establishment of pronominal reference in the case of languages with relative clauses containing null pronominals (Chinese, Japanese, Korean).

A second factor, not necessarily incompatible with the previous one, involves a typological trait of Basque language that is not shared by any of the other languages in these studies: ergativity (Dixon, 1994. Marantz, 1984). In ergative languages like Basque there is a single (typically unmarked) morphological class that includes both intransitive/thematic subjects and transitive objects such as "Ikaslea" in (7a), (7b), and another morphological class, typically marked, consisting of transitive/agentive subjects such as "irakasleak" in (7b). In the case of Basque, this class displays an ergative case ending (*-k*, glossed erg):

(7a)	ikasle-a etorri da student-sg arrived is "the student arrived"	intransitive/thematic subject
(7b)	irakasle-a-k ikasle-a ikusi du teacher-sg-erg student-sg seen has "the teacher saw the student"	transitive subject

All ergative languages share this property, and in fact this is what ergativity is: grouping in one unmarked class objects and intransitive/thematic subjects, and in another, marked class, transitive/agentive subjects. Thus, the central difference between nominative/accusative languages like English, and ergative/absolutive languages like Basque is how they mark this set of arguments: nominative languages group together transitive/intransitive subjects and separate objects, while ergative languages group together intransitive subjects and objects, and separate transitive subjects, as illustrated in Fig. 6.

This entails that, given an ergative language and a transitive clause, it is the object that is morphologically equivalent to the subject of an intransitive clause. On the contrary, given a nominative–accusative language and a transitive clause, it is the subject that is equivalent to the subject of an intransitive clause. As shown in Fig. 6, the shaded classes are morphologically marked, -and the clear classes are unmarked.

If morphological unmarkedness provides a processing advantage in language, then different patterns of complexity would arise in each type of language: nominative–accusative languages will display a subject advantage, but ergative languages will display an object (and intransitive subject) advantage in gap-filler dependencies. This typological trait might thus explain why an object relative clause preference has been found in Basque, where objects are the unmarked form, contrary to both head-initial and head-final nominative–accusative languages like English,

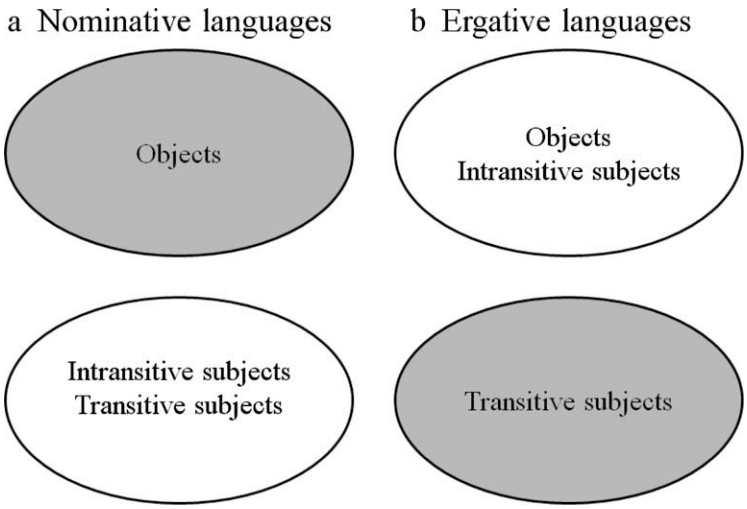


Fig. 6. Representation of argument marking for nominative/accusative languages like English (a) and ergative/absolutive languages like Basque (b).

Spanish, Japanese or Korean, where objects are the marked class.

7. Conclusions

To conclude, our experiments have shown that Basque SR are harder to process than OR, in both reading times and ERPs at least with this particular configuration. Our ERP and reading time data seem more consistent with the integration cost hypothesis (see Hsiao & Gibson, 2003) than with the universal structural distance account. However, the integration cost hypothesis predicts an OR advantage for all head-final languages with pronominal relative clauses, which appears not to be the case given the results for Korean and Japanese. We suggest that the underlying processing mechanism at play is probably not dependent on notions like subject and object, but rather on argument-marking classes, so that nominative-gap RCs are easier than accusative-gap RCs, whereas in ergative languages, it is the absolutive-gap RCs that have a processing advantage over ergative-gap RCs. In generative linguistic theory it is assumed since Chomsky (1965) that notions like subject and object are derivative, not primitives of syntactic structure; it is therefore not unlikely that language processing mechanisms may turn out not be based on such notions either, but rather on formal, morphological aspects that can vary across grammar types, as does in the nominative/ergative divide.

Acknowledgment

This research was partially supported by Grants from the Spanish Ministry of Education and Science (SEJ2006-09238, SEJ2007-60751/PSIC, CONSOLIDER INGENIO 2010 CSD2008-00048 and CSD2007-00012), from the Basque Government (BFI05.310), and from the University of the Basque Country (UPV/EHU GIU06/52). Correspondence concerning this paper should be sent to Manuel Carreiras. Basque Center on Cognition, Brain and Language. Paseo Mikeletegi, 53, 20009 Donostia-San Sebastián, Spain.

References

Aduriz, I., Aranzabe, M. J., Arriola, J. M., Atutxa, A., Díaz de Ilarraza, A., Ezeiza, N., et al. (2006). Methodology and steps towards the construction of EPEC, a corpus of written Basque tagged at morphological and syntactic levels for the automatic processing. In A. Wilson, P. Rayson, & D. Archer (Eds.), *Corpus linguistics around the world*. *Language and computers* (Vol. 56. 90-420-1836-4, pp. 1–15). The Netherlands: Rodopi.

Artiagoitia, X. (1992). Why Basque doesn't relativize everything? In Lakarra & Ortiz de Urbina (Eds.), *Syntactic Theory and Basque Syntax* (Vol. 27, pp. 11–35) San Sebastian: "Julio de Urquijo" International Basque Philology Seminar.

Artiagoitia, X. (1990). *A Unified Account of Relativization in Basque*, U. of Washington, MA thesis.

Baker, M. (2001). *The atoms of language*. New York: Basic Books.

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in spanish: An ERP comparison. *Journal of Cognitive Neuroscience, 17*(1), 137–153.

Betancort, M., Carreiras, M., & Sturt, P. (2009). The processing of subject and object relative clause in Spanish: An eye-tracking study. *Quarterly Journal of Experimental Psychology, 62*, 1915–1929.

Bever, T. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279–352). New York: John Wiley & Sons, Inc..

Boland, J. E. (1997). The relationship between syntactic and semantic processes in sentence comprehension. *Language and Cognitive Processes, 12*, 423–484.

Caplan, D., Vijayan, S., Kuperberg, G., West, C., Waters, G., Greve, D., et al. (2002). Vascular responses to syntactic processing: Event-related fMRI study of relative clauses. *Human Brain Mapping, 15*, 26–38.

Carreiras, M., Salillas, E., & Barber, H. (2004). Event related potentials elicited during parsing of ambiguous relative clauses in Spanish. *Cognitive Brain Research, 20*, 98–105.

Chien-Jer, C.L. & Bever, T. (2006). Subject Preference in the Processing of relative Clauses in Chinese. In D. Baumer, D. Montero, and M. Scanlon (Eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, Somerville, MA: Cascadilla Proceedings Project, 254–260.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky, 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT Press.

Chomsky, N. (1986). *Knowledge of language: Its nature, origins and use*. New York: Praeger.

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

Cohen, L., & Mehler, J. (1996). Click monitoring revisited: An on-line study of sentence comprehension. *Memory and Cognition, 24*, 94–102.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes, 13*(1), 21–58.

Dixon, R. M. W. (1994). *Ergativity, Cambridge studies in linguistics*. Cambridge: Cambridge University Press. 69.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language, 67*, 547–619.

Ford, M. (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior, 22*, 203–218.

Frauenfelder, U., Segui, J., & Mehler, J. (1980). Monitoring around the relative clause. *Journal of Verbal Learning and Verbal Behavior, 19*, 328–337.

Frazier, L. (1987). Syntactic processing: Evidence from Dutch. *Natural Language & Linguistics Theory, 5*, 519–559.

Frazier, L., & Flores D'Arcais, G. (1989). Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language, 28*, 331–344.

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

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

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

Gennari, S. P., & MacDonald, M. C. (2008). Semantic indeterminacy in object relative clauses. *Journal of Memory and Language, 58*, 161–187.

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

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. Marantz, & W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). Cambridge, MA: MIT Press.

Gibson, E., Hickok, G., & Schutze, C. (1994). Processing empty categories in a parallel parsing framework. *Journal of Psycholinguistic Research, 23*, 381–405.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning Memory and Cognition, 27*, 1411–1423.

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language, 51*, 97–114.

Greenhouse, S., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95–111.

Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science, 29*, 261–290.

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

Hagoort, P., Brown, C., & Osterhout, L. (1999). The neurocognition of syntactic processing. In C. Brown & P. Hagoort (Eds.), *Neurocognition of language* (pp. 273–316). Oxford, UK: Oxford Univ. Press.

Hale, J. (2003). The information conveyed by words in sentences. *Journal of Psycholinguistic Research, 23*, 101–123.

Holmes, V. M., & O'Regan, J. K. (1981). Eye fixation patterns during the reading of relative-clause sentences. *Journal of Verbal Learning and Verbal Behavior, 20*, 417–430.

Hsiao, F., & Gibson, E. (2003). Processing relative clause in Chinese. *Cognition, 90*, 3–27.

Ishizuka, T. (2005). Processing relative clauses in Japanese. In Okabe and Nielsen (Eds.), *UCLA Working Papers in Linguistics, no.13, Papers in Psycholinguistics* (Vol. 2, pp. 135–157.

Ishizuka, T., Nakatani, K., & Gibson, E. (2006). Processing Japanese relative clauses in context. In *Paper presented at The 19th annual CUNY conference on human sentence processing*. New York, NY: CUNY Graduate Center.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329–354.

Kaan, E., & Swaab, T. (2003). Electrophysiological evidence for serial sentence processing: A comparison between non-preferred and ungrammatical continuations. *Cognitive Brain Research, 17*, 621–635.

Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry, 8*, 63–99.

Keenan, E. L., & Hawkins, S. (1987). The psychological validity of the accessibility hierarchy. In E. Keenan (Ed.), *Universal grammar: 15 essays* (pp. 60–85). London: Routledge.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*, 580–602.

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

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

Koopman, H., & Sportiche, D. (1991). The position of subjects. *Lingua, 85*, 211–259.

Kuo, K., & Vasishth, S. (submitted for publication). *Processing Chinese relative clauses: Evidence for the universal subject preference*.

Kutas, M., Van Petten, C., & Kluender, R. (2006). Psycholinguistics electrified II: 1995–2005. In M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 659–724). New York: Elsevier Press.

Kwon, N., Lee, Y., Gordon, P., Kluender, R., & Polinsky, R. (submitted for publication). *Cognitive and linguistic determinants of the subject–object asymmetry: An eye-tracking study of pre-nominal relative clauses in Korean*.

Kwon, N., Polinsky, M., & Kluender, R. (2006). Subject preference in Korean. In D. Baumer, D. Montero, & M. Scanlon (Eds.), *Proceedings of the 25th west coast conference on formal linguistics* (pp. 1–14). Somerville, MA: Cascadilla Proceedings Project.

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

Lasnik, H., & Uriagereka, J. (1988). *A course on GB syntax. Lectures on binding and empty categories*. Cambridge: MIT Press.

Lin, C.-J., & Bever, T.G. (2006). Chinese is no exception: Universal subject preference of relative clause processing. In *Paper presented at the The 19th annual CUNY conference on human sentence processing*. New York, NY: CUNY Graduate Center.

Lin & Garnsey (submitted for publication). *Animacy and the resolution of temporary ambiguity in relative clause comprehension in Mandarin*.

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes, 9*, 157–201.

MacDonald, M. C., & Christiansen, M. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1999). *Psychological Review, 109*, 35–54.

MacDonald, M., Pearlmutter, N., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*, 676–703.

MacWhinney, B. (1977). Starting points. *Language, 53*, 152–168.

MacWhinney, B. (1982). Basic syntactic processes. In S. Kuczaj (Ed.), *Syntax and semantics (1). Language acquisition*. Hillsdale, NJ: Erlbaum.

MacWhinney, B., & Pleh, C. (1988). The processing of restrictive relative clauses in Hungarian. *Cognition, 29*, 95–141.

Mak, P., Vonk, W., & Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language, 47*, 50–68.

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

Marantz, A. (1984). *On the nature of grammatical relations*. Cambridge, MA: MIT Press.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language, 38*, 283–312.

Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory and Cognition, 23*, 477–494.

Mitchell, D. C. (2004). On-line methods in language processing: Introduction and historical review. In M. Carreiras & C. Clifton, Jr. (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP, and beyond*. Psychology Press.

Mitchell, D. C., Cuetos, F., Corley, M. M. B., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research, 24*, 469–488.

O'Grady, W. (1997). *Syntactic development*. Chicago, IL: Chicago University Press.

O'Grady, W., Miseon, L., & Miho, C. (2003). A subject–object asymmetry in the acquisition of relative clauses in Korean as a second language. *Studies in Second Language Acquisition, 25*, 433–448.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*, 785–804.

Osterhout, L., McLaughlin, J., & Bersick, M. (1997). Event-related brain potentials and human language. *Trends in Cognitive Sciences, 1*, 203–209.

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34*, 739–773.

Oyharçabal, B. (1987). *Etude descriptive de constructions complexes en basque: propositions relatives, temporelles, conditionnelles et concessives,* Doctoral dissertation, Université de Paris 7, Microfilm: Atelier de publication des thèses de l'Université de Lille 3.

Pickering, M. J. (1994). Processing local and unbounded dependencies: A unified account. *Journal of Psycholinguistic Research, 23*, 323–352.

Pickering, M., & Barry, G. (1991). Sentence processing without empty categories. *Language and Cognitive Processes, 3*, 229–259.

Pickering, M., & Barry, G. D. (1993). Dependency categorial grammar and coordination. *Linguistics, 31*, 855–902.

Pollock, J. Y. (1989). Verb movement, UG, and the structure of IP. *Linguistic Inquiry, 20*, 365–424.

Sauerland, U., & Gibson, E. (1998). How to predict the relative clause attachment preference. In *Paper presented at The 11th annual CUNY conference on human sentence processing*. New Brunswick, NJ: Rutgers University.

Schriefers, H., Friederici, A. D., & Kühn, K. (1995). The processing of locally ambiguous relative clauses in German. *Journal of Memory and Language., 34*, 499–520.

Spivey-Knowlton, M., & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition, 55*, 227–267.

Stowe, L. (1989). Thematic structures in language processing. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing*. Dordrecht: Kluwer Academic Publishers.

Tabor, W., Juliano, C., & Tanenhaus, M. K. (1997). Parsing in a dynamical system: An attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language & Cognitive Processes, 12*, 211–272.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language, 47*, 69–70.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33*, 285–318.

Trueswell, J. C., Tanenhaus, M., & Kello (1993). Verb specific constraints in sentence processing: Separating effects of lexical preference from garden-path. *Journal of Experimental Psychology: Learning, Memory and Cognition, 19*, 528–553.

Ueno, M., & Garnsey, S. (2008). An ERP study of the processing of subject and object relative clauses in Japanese. *Language & Cognitive Processes, 23*, 646–688.

Van Berkum, J. J. A., Hagoort, P., & Brown, C. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11*, 657–671.

Wanner, E., & Maratsos, M. (1978). An ATN approach in comprehension. In M. Halle, J. Bresnan, & G. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 119–161). Cambridge, MA: MIT Press.

Weckerly, J., & Kutas, M. (1999). An electrophysiological analysis of animacy effects in the processing of object relative sentences. *Psychophysiology, 36*(5), 559–570.