To appear in Peter de Swart and Monique Lamers (eds.). *Case, Word Order, and Prominence. Interacting Cues in Language Production and Comprehension*. Springer.

10. Processing verb medial word orders in a verb final language

Kepa Erdocia[1] ,Itziar Laka[1] ,Antoni Rodriguez-Fornells[2, 3]

[1] University of the Basque Country UPV/EHU

[2] Dept. Ciencies Fisiològiques II - IDIBELL, University of Barcelona, Barcelona, Spain

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

Abstract We compared the processing of verb-medial sentences in Basque. Syntactic analysis claims that all word orders other than SOV are derived in this language; therefore, verb medial sentences are expected to show signs of syntactic displacement and be equally complex to process. A self-paced reading task measured processing time for each constituent of the sentences in order to analyze the processing of verb-medial SVO and OVS structures. Sentences with verb-medial word orders did not show any significant reading time difference. This result suggests that both SVO and OVS sentences are indeed syntactically derived from the canonical SOV word order, shown in previous work to be easiest computationally.

10.1. INTRODUCTION

One important question in the history of Linguistics is whether language variation is unbounded, or whether there are universal principles that constrain the space for variation (Robins 1967). Joseph Greenberg's *Universals of Language* (1963) and Chomsky's *Syntactic Structures* (1957) constitute landmarks that changed the focus and framework for discussion. Greenberg's work showed there are significant generalizations emerging from various aspects of linguistic variation. Chomsky's approach successfully argued that the goal of linguistics is to find "the fundamental underlying properties of successful grammars. The ultimate outcome of these investigations should be a theory of linguistic structure in which the descriptive devices utilized in particular grammars are presented and studied abstractly, with no specific reference to particular languages." (Chomsky 1957: 11). Since then, it is established that variation across languages is limited to a rather narrow range, and human languages are organized following the same basic patterns and mechanisms. Greenberg was the founder of modern linguistic typology; the results of his surveys showed that many language features are related, and discarded the idea of limitless variation; some decades later, in the

eighties, the *Principles and Parameters* model articulated by Chomsky (1981) claimed that those patterns of variation can be derived from a finite set of linguistic parameters. Languages can thus be classified by these structural features, or *parameters*, into a few language types (Baker 2001).

Here, we focus on the implications that arise from variations in the relative basic order of the S(ubject), O(bject), and V(erb) in a sentence. What Greenberg found in his language-survey is that apparently unrelated syntactic properties are strongly correlated: if a language has SOV as its canonical constituent order in declarative sentences, then this language will tend to have postpositions, relative clauses will appear before nouns, adverbs before verbs or adjectives, and main verbs before auxiliary verbs. Following Greenberg's implicational universals, Basque language has been argued to be of the SOV type (de Rijk, 1969). In Basque, declarative sentences that initiate a discourse (that is, sentences where all information is new) display SOV order. Following the greenbergian correlations, Basque has postpositions (example 1), in declarative sentences adverbs appear to the left of verbs and adjectives (examples 3 and 4), and auxiliary verbs appear following the main verb (example 5).

(1) gizon-A-REKIN
man-det-with
'WITH THE man'
(2) [gizon-a ikusi duEN] emakume-a
man-det seen has-REL woman-det
"the woman that saw the man"
(3) Mikel BERANDU etorri da
Mikel LATE arrived is
'Mikel arrived LATE'
(4) Zu-re auto-a OSO polita da
You-GEN car-det VERY beautiful is
'Your car is VERY beautiful'
(5) Mikel-ek Ana ikusi DU
Mikel-ERG Ana seen HAS
'Mikel HAS seen Ana'

10.1.1. Word Order

Basque is a free word order language. Constituents can appear in almost any order. In example (6) which is adapted from Ortiz de Urbina (2003: 448), all sentences generated combining the constituents in all orders (in square brackets) are grammatical:

(6) a. PP-S-IO-O-V

[$_{PP}$Afaldu ondoren] [$_{SUBJ}$Mikel-ek] [$_{IO}$Ana-ri] [$_{OBJ}$gerriko berria] [$_{V+AUX}$oparitu dio]
[diner after] [Mikel-ERG] [Ana-DAT] [belt new-det] [given aux]
'*After dinner, Mikel has given the new belt to Ana*'

b. O-PP-IO-S-V
[$_{OBJ}$Gerriko berria] [$_{PP}$afaldu ondoren] [$_{IO}$Anari] [$_{SUBJ}$Mikelek] [$_{V+AUX}$oparitu dio]

c. S-O-PP-V-IO
[$_{SUBJ}$Mikelek] [$_{OBJ}$gerriko berria] [$_{PP}$afaldu ondoren] [$_{V+AUX}$oparitu dio] [$_{IO}$Anari]

d. IO-V-O-PP-S
[$_{IO}$Anari] [$_{V+AUX}$oparitu dio] [$_{OBJ}$gerriko berria] [$_{PP}$afaldu ondoren] [$_{SUBJ}$Mikelek]

e. ...

Based on this constituent-order freedom in the sentence, some linguists (Rebuschi 1987) have argued that sentence structure in Basque is flat and has no hierarchical arrangement (that is, it is non-configurational). However, most authors agree that sentence structure is hierarchical in this language (De Rijk 1969, 2007; Ortiz de Urbina, 1989; Laka, 1990; Artiagoitia, 1995; Fernandez, 1998; Elordieta 2001, among many others). In particular, evidence from linguistic theory shows that the Subject constituent is higher than the Object constituent. In other words, there is widespread agreement among linguists that (6a) is the underived word order of a sentence, and that other word-orders are derived from this one via syntactic movement/displacement of constituents.

10.1.2. Syntactic complexity

According to syntactic studies of Basque grammar (De Rijk 1969, Ortiz de Urbina 1989, Laka 1990, Artiagoitia 1995, Fernandez 1998, Elordieta 2001, among others), SOV sentences reflect the canonical word order, and all other word orders (OSV, SVO and OVS) are syntactically derived. Assuming this to be the correct analysis of Basque sentence grammar, we hypothesized that SOV would be easiest to process, and all other word orders (OSV, SVO and OVS) would show a processing cost due to the extra syntactic operations involved.

As shown in the (somewhat simplified) syntactic representation in Figure 10.1, the syntactic structure of the OSV word order can be derived from the canonical SOV by means of displacing the object over the subject.

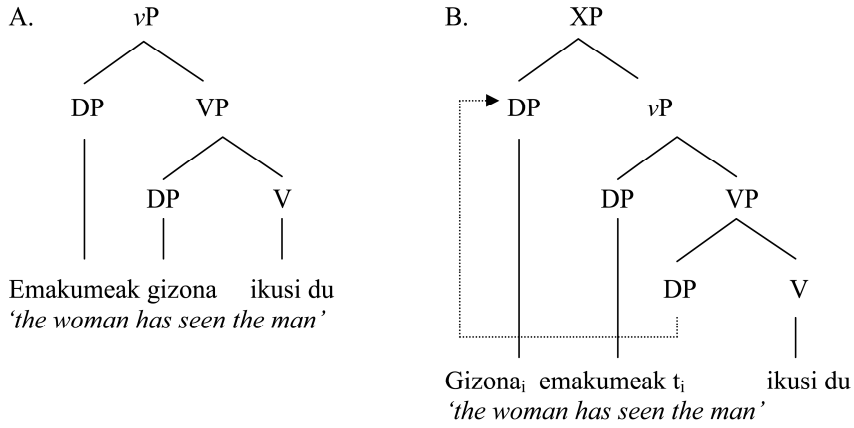

Fig. 10.1: *In these syntactic representations the* v*P represents the transitive sentence where the Determiner Phrases (DP) are the subject and the object of the verb (V). A: Syntactic structure of the canonical SOV word order in Basque. The object (gizona 'man') is generated in the Verb Phrase (VP), preceding the verb (V). B: Starting from the basic SOV order, OSV is generated displacing the object above the subject to the specifier position of a higher projection (XP).*

Recently obtained psycholinguistic and neurolinguistic evidence supports the linguists' claim that Basque language's SOV word order is syntactically simpler than OSV word order (Erdocia et al., 2009): reading times showed that canonical SOV is processed faster than OSV word order by native speakers of this language. In the OSV condition, reading the Subject after the Object required extra processing effort comparing with reading the Subject before the Object in the SOV condition. The derived OSV word order triggered syntactic reanalysis in second DP[1] position probably due to the fact that participants processed the first DP as the Subject of an intransitive sentence (see Fig. 10.2).

[1] Throughout the paper, we will use the label DP (Determiner Phrase), which is more accurate than NP (Noun Phrase) for the syntactic label of the Subject and Object constituents involved.

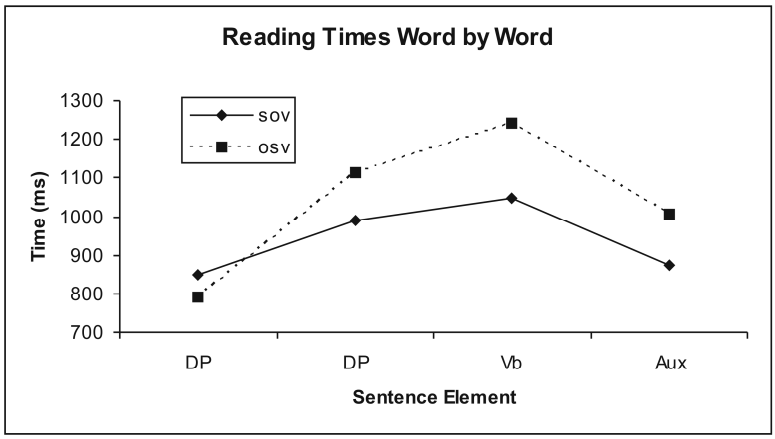

Fig. 10.2: *Word-by-word mean reading times of canonical SOV and derived OSV word orders. The y-axis features reaction times in milliseconds; in x-axis the different constituents of each sentence are depicted (DP = Determiner Phrase; Vb = Verb; Aux = Auxiliary verb) (adapted from Erdocia et al., 2009).*

In addition to reading times, ERPs provided evidence favoring the interpretation that OSV is syntactically derived: OSV sentences showed increasing negativities in both Subject and Object DP positions (Fig. 10.3). These negativities at DP positions suggested that Subjects and Objects were processed differently. A P600 effect was observed in the Verb position when compared to the canonical SOV word order; this P600 effect reflects the higher processing cost of the derived OSV word order.

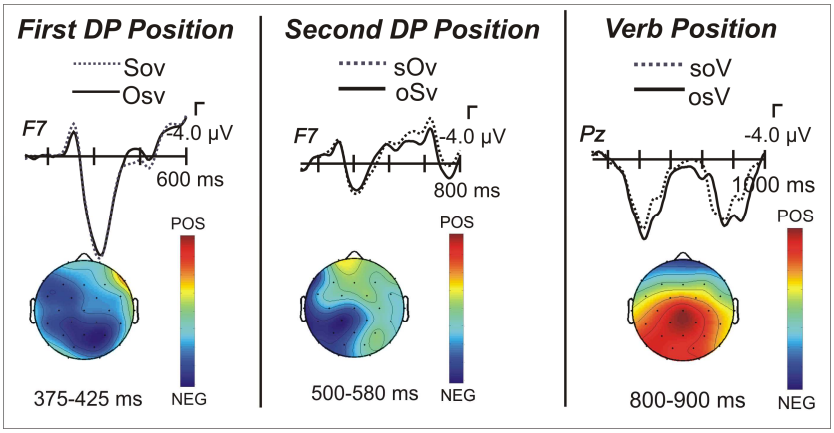

Fig. 10.3: *ERP comparison of SOV and OSV word orders in Basque. In first and second DP positions, the negativities indicated that Subjects and Objects were differently processed. At verb position, the late P600 effect represents the higher processing cost of derived words orders (adapted from Erdocia et al., 2009).*

Here we report results obtained from experiments designed to investigate how verb-medial sentences (SVO/OVS) are processed by native speakers of Basque. Verb-medial word orders are taken to be derived from canonical SOV structures (Etxepare and Ortiz de Urbina 2003). The syntactic representations below (Fig. 10.4) illustrate schematically how the verb-medial structures are derived from SOV structure.

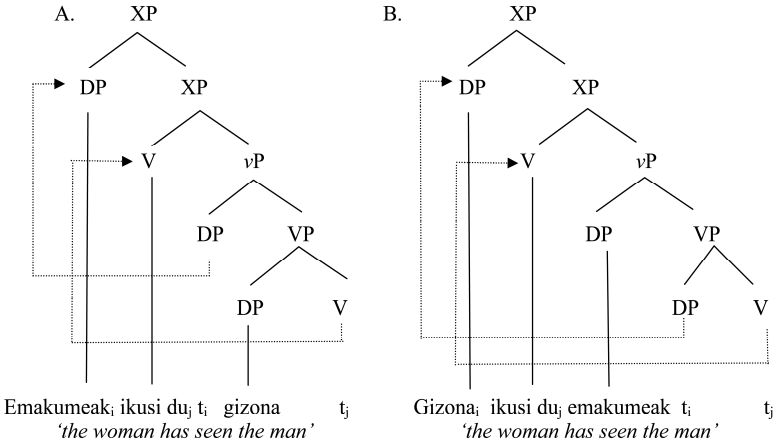

Fig. 10.4: *A: Syntactic representation of the derived SVO word order in Basque. The subject (emakumeak 'woman') and the verb (ikusi du 'has seen') are displaced above the object. B: The OVS is derived displacing the object and verb above the subject.*

10.1.3. Pro-Drop and Ergativity in Basque

Basque has a pluripersonal verb-agreement system, as a consequence of which it is a three-way pro-drop language (Ortiz de Urbina 1989, Laka 1993, 1996); this means that subjects, objects and datives can be phonologically unrealized:

(7) a. pro_{SUBJ} pro_{IO} pro_{OBJ} eman d-i-zu-t
(I) (you) (it) given it-root-you-me
'*I have given it to you*'
b. pro_{SUBJ} pro_{OBJ} ikusi na-u-zu
(you) (me) seen me-root-you
'*you have seen me*'

Basque is also an ergative language (Levin 1983, Laka 2006): subjects of intransitive clauses and objects of transitive clauses are morphologically identical, and bear no case ending (8a, b), while subjects of transitive clauses are morphologically marked by an ergative case marker (8c):

(8) a. **Gizon-a** etorri da (subject of intransitive sentence)
man-det arrived is
'*The man arrived*'
b. Emakume-a-k **gizon-a** ikusi du (object of transitive sentence)
woman-det-ERG man-det seen has
'*The woman has seen the man*'
c. **Gizon-a-k** emakume-a ikusi du (subject of transitive sentence)
man-det-ERG woman-det seen has
'The man saw the woman'

Given this combination of grammatical features, if the DP constituent *gizon-a* "the man" is encountered at the beginning of an utterance, the following possibilities arise:

(i) it is the Subject of an SV intransitive sentence like (8a);
(ii) it is the Object of an OSV sentence like (8b);
(iii) it is the Object of an OV transitive sentence where S has been omitted as in (7a,b).

10.1.4. Ambiguities

We also investigated how syntactically ambiguous sequences are processed. Basque has a morphologically ambiguous form shown in (9), where both subject and object constituents present the same ending *–ak*.

(9) Emakume-**a-k** gizon-**ak** ikusi d-it-u
woman-det_{sg}-ERG man-det_{pl} seen 3_{pl}-has-3_{sg}
`The woman has seen the men'

In (9), the singular transitive subject of the sentence *emakumeak* "the woman (erg)" carries two morphemes: the singular determiner morpheme *–a* we have seen in previous examples (8a, b), followed by the ergative case marker *–k* also illustrated before (8c). The plural object of the sentence *gizonak* "the men" carries only one morpheme, the plural determiner *–ak*, which happens to be homophonous with the subject's morphological sequence. Moreover, given that in Basque objects can appear at the beginning of the sentence, it turns out that the sentence in (9), if prosodic information is not provided and it is only presented visually, could also be interpreted as an OSV sentence like (10):

(10) Emakume-**ak** gizon-**a-k** ikusi d-it-u
woman- det_{pl} man- det_{sg}-ERG seen 3_{pl}-has-3_{sg}
`the man has seen the women'

By measuring the reading times of fully ambiguous verb final declarative sentences (e.g., sentences like 9 and 10) we have shown in a previous study (Erdocia et al., 2009) that they are straightforwardly processed as SOV canonical sentences by native speakers. We interpret these results as showing that the first DP constituent is processed as a singular ergative subject (NP-*a*-*k*), and hence the second DP constituent is processed as plural object (NP-*ak*). Hence, verb final word orders seem to reveal a strong *subject first* processing strategy in Basque (Bates et al., 1988, see also Kretzschmar et al. and Lamers in this volume).

Furthermore, as shown in (11), a sentence can be ambiguous with respect to its initial constituent, and be morphosyntactically disambiguated by a subsequent constituent. That is, we can start a sentence marking the first DP constituent with the ambiguous *–ak* ending, so that temporarily it could either be an ergative singular Subject or a plural Object, but if the second constituent is morphologically unambiguous, the structure is instantly disambiguated.

(11)	Emakume-**ak**	gizon-**ek**	ikusi	d-it-uzte
	woman- det_{pl}	man-det_{pl}-ERG	seen	3-pl-has-3_{pl}
	`the men saw the women'			

Example (11) starts ambiguously (the first constituent could be the Subject or the Object of the sentence) but it is disambiguated at the second constituent position by means of the unambiguous *–ek* plural ergative subject marker. This plural ergative *–ek* ending results from the merging of the plural determiner *–ak* and the ergative case marker *–k* via morphophonological processes of epenthesis and deletion that we do not detail here (see de Rijk 2007). Thus, once the second constituent is processed, the first constituent should necessarily be interpreted as the Object of the sentence. In our previous study, we showed that these temporally ambiguous sequences (e.g., sentence type 11) are disambiguated syntactically at the second constituent position (Erdocia et al., 2009). Comparing the canonical SOV sentences and these temporarily ambiguous OSV sentences, the reading times increased significantly at sentence second position of the temporally ambiguous sentences (Fig. 10.5).

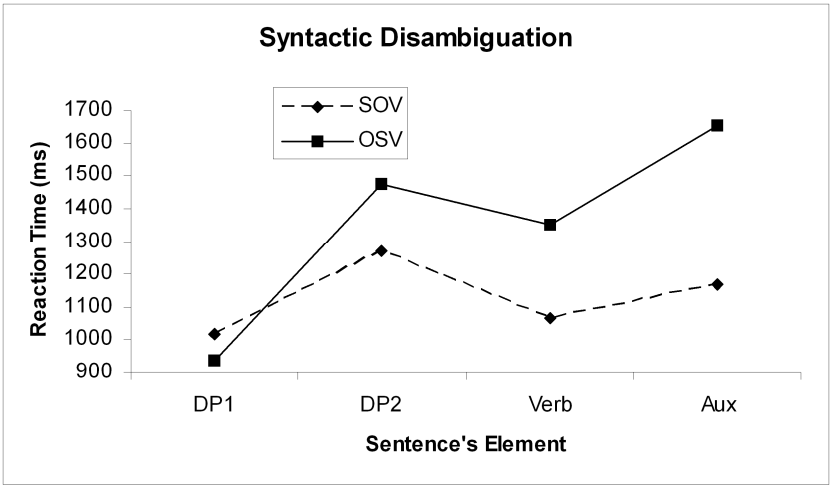

Fig. 10.5: *Syntactic Disambiguation effect at second DP position of OSV derived sentences (adapted from Erdocia et al., 2009).*

Thus, the result of temporally ambiguous sentences suggests that participants initially processed the first ambiguous constituent as the subject of the sentence, but upon finding the unambiguous plural subject at sentence second position, they were forced to reanalyze their initial processing hypothesis, changing the *subject-before-object* interpretation to the derived *object-before-subject* interpretation. In an ERP experiment, (Erdocia et al., 2009, exp. 3) we also observed a specific modulation of a negative frontal-posterior component at the disambiguation point of temporally ambiguous sentences which were disambiguated by means of the world knowledge at the verb constituent. The frontal-posterior negativity spread from 500 until 1000 ms after the disambiguation point. We interpreted this modulation as an index related to working memory costs induced by syntactic reinterpretation analysis (changing the default SOV interpretation to OSV in ambiguous sentences).

In the present study we took advantage of these morphologically ambiguous *–ak* markers to explore the syntactic disambiguation effect in verb medial structures: SVO and OVS. Furthermore, we explored whether these syntactically derived (non-canonical) structures are equally or differently processed in terms of reading times.

10.2. COMPARING DERIVED SVO AND OVS WORD ORDERS

The aim of the present study was to measure the syntactic complexity and the ambiguity resolution processes that word order variation generates. By using *self-paced reading* (Just et al. 1982) and a *comprehension task*, we studied the processing of sentences with SVO and OVS word orders in Basque, in order to measure differences between the *subject-before-object* and *object-before-subject* effects.

10.2.1. Material

The experimental material consisted of 100 sentences like (12a). Starting from this type of sentence, we generated the rest of conditions by means of changing the order and the number of constituents. Thus we have 100 sentences of singular SOV structure (12a), 100 sentences of singular OSV (12b), 100 sentences of ambiguous structure (12c), 100 sentences of plural SOV structure (12d), and finally 100 sentences of plural OSV (12e). Altogether there were 500 stimuli.

(12) a. Gizon-a-k ikusi du emakume-a [S_{SING}-V-O_{SING}]
Man-det_{sg}-ERG see has woman-det_{sg}
'The man has seen the woman'

b. Emakume-a ikusi du gizon-a-k [O_{SING}-V-S_{SING}]
Woman-det_{sg} see has man-det_{sg}-ERG
'The man has seen the woman'

c. Gizon-ak ikusi ditu emakume-ak[AMB-V-AMB]
man-AMB see has woman-AMB
'The man has seen the women' or
'The woman has seen the men'

d. Gizon-ek ikusi dituzte emakume-ak [S_{PL}-V-O_{PL}]
men-det_{pl}-ERG see have women-det_{pl}
'The men have seen the women'

e. Emakume-ak ikusi dituzte gizon-ek [O_{PL}-V-S_{PL}]
women-det_{pl} see have men-det_{pl}-ERG
'The men have seen the women'

Recall that the constituents carrying the morpheme *-ak* are ambiguous. Therefore, sentences (12a), (12c) and (12e) are ambiguous at first constituent position. Sentences (12a) and (12e) are disambiguated at auxiliary position. In the case of (12a) the resulting structure is SVO and in the case of (12e) the resulting structure

is OVS. Only sentences like (12c) are fully ambiguous because both constituents carry the *-ak* ending[2].

Stimuli were divided in five lists. Sentences were randomized (Latin-square), and one version of each item was assigned to one of the five lists. This method allowed every participant to read only once each version of a sentence. The five lists were balanced across participants ensuring that the material was correctly rotated across conditions and participants. This ensured that each version of each sentence was read equally number of times across participants. The experimental items were presented along with 16 practice trials (2 sentences of each condition and 6 fillers) to verify that participants understood the instructions of the experiment.

We created 100 filler sentences which were also presented in the experiment. Half of the fillers were intransitive sentences with the absolutive subject realized (see example 13). The remaining 50 fillers were ditransitive sentences where the ergative subject was dropped. The absolutive direct object and the dative indirect object were realized in the sentence and they varied in the position they occupied in the sentences (see examples 14a and 14b)

(13) Mutiko-a aizto-a-z ebaki da
Boy-det$_{sg}$ knife-det$_{sg}$-with cut is
'The boy has cut himself with the knife'

(14) a. *pro* etorkin-a-ri hil d-i-o-te ume-a
(they-ERG) immigrant-det$_{sg}$-DAT kill 3$_{sg}$-have-3$_{DAT}$-3$_{pl}$ child-det$_{sg}$
'They have killed the immigrant's child'

b. *pro* idazle-a aurkeztu d-i-o-te itzultzaile-a-ri
(they-ERG) writer-det$_{sg}$ introduce 3$_{sg}$-have-3$_{DAT}$-3$_{pl}$ translator-det$_{sg}$-DAT
'They have introduced the writer to the translator'

The same fillers were used in the five lists. Like in the experimental conditions, fillers were four words long.

10.2.2. Method

The experimental data were obtained by presenting the materials to the participants using Self Paced Reading Moving Window (Just et al. 1982) and performing the Comprehension Task. We measured the reading times of the participants, which is the time required for reading each word in every sentence. Then we added the reading time of each word, in order to obtain the reading time of the whole sentence. We could determine whether the participants understood

[2] Although conditions (12a), (12c) and (12e) start ambiguously we decided to label only the constituents of the fully ambiguous condition (12c) with the *AMB* gloss.

the sentences correctly through the comprehension task, and we also measured the reaction time of the comprehension task itself.

Different methods could be chosen for stimuli presentation using the self-paced reading technique (Just et al. 1982; Kennedy and Murray 1984). Due to the agglutinative morphology of Basque language we decided to present the sentences word by word. In this manner participants read the words as they would normally read them in a written text. The moving window technique used here (Just et al. 1982) requires that sentences are presented word by word, and once a word has been read it turned into a string of asterisks. When participants finished reading the sentence this way, a comprehension question appeared on the screen. In order to answer the question of the comprehension task, participants had to press the button corresponding to 1 and 2 in the keyboard. Participants did not receive feedback about their answers. The amount of time the participant spent reading each word was recorded as the time between key-presses.

The comprehension task ensured that the participants had understood the sentences they read. In the case of fully ambiguous conditions, the comprehension task allowed us to know the preferred interpretation of the participants when confronting fully ambiguous sentences. That is, the comprehension task showed whether a participant interpreted an ambiguous chain as SVO or as OVS. The task consisted in a yes-or-no question after each sentence. The answer to half of the questions of each word order was "yes" and the other half was "no". Example (15) would be the question corresponding to the test item of (12a); in this case the answer would be positive:

(15)	egi-a	al	da	emakume	bat-ek	ikusi du-ela	gizon bat?
	true	Q[3]	is	woman	one-erg	seen has-comp	man one
	'is it true that one woman has seen a man?'						

For each variable (DP1, DP2, V, Aux, Total Reading Times, Reaction Time and Amount of Errors in the CT), an omnibus Anova analysis was carried out, introducing Condition as a general factor (5 levels). When this effect was significant, we computed pairwise comparisons between the different conditions. We used Bonferroni correction in order to prevent false positive results (for 10 pairwise comparisons in each variable, we only report those effects with a $p < 0.005$).

10.2.3. Participants

Twenty nine native speakers of Basque participated in the experiment as volunteers, all of them undergraduate students of the Faculty of Letters (University of the Basque Country UPV/EHU) All participants who made more

[3] The interrogative particle *al* is used in yes-no questions.

than 7 mistakes out of 20 sentences of each condition were rejected from statistical analysis. However, in the case of ambiguous sentences we included the participants who made more than 7 mistakes for statistical analysis since in any case their answers could not be taken as incorrect. All in all, we analyzed data from 24 subjects (2 men and 22 women[4]; mean age 20, SD ± 3.21).

The participants filled a Basque-Spanish bilingualism questionnaire (adapted from Weber-Fox and Neville 1996). The results of the questionnaire showed that participants use the Basque language more often than Spanish[5]. More specifically the results for Basque obtained 1.9 ± 0.3 (SD) points in a 1 to 7 point scale where 1 corresponded only to Basque use and 7 only Spanish). Subjects had to rate themselves in comprehension, in reading, in speaking and in writing. In a 4-point scale where 4 was 'very good' and 1 was 'bad', the mean values for the present sample were 3.94 ± 0.2 for Basque and 3.5 ± 0.6 for Spanish. They were also asked in which language they felt more comfortable or preferred (in general); six participants answered they felt equally comfortable in both languages, one preferred Spanish, and the rest of the participants felt more comfortable in Basque. In sum, the present sample is very fluent in Basque and is the language which is usually used in their environment.

10.2.4. Recording

Participants were instructed to read at a normal, comfortable pace and in a way that would enable them to understand the sentences. Sentences were presented using a standard Personal Computer and running the EXPE6 (Pallier et al. 1997) program. The computer recorded the time interval since a word was first displayed until the next press of the space bar. The choice program recorded the reaction times and the answers of the participants. More specifically the program recorded: (i) the required time to read each word of the sentence, (ii) the required time to read and answer the question of the comprehension task, and (iii) whether the answer to the question was correct or not.

---

[4] In previous literature, no differences between males and females have been found in language processing. Given this, we did not particularly seek to keep equal numbers of males and females in the groups.

[5] One of the participants, who was from Hendaye (a town in the French side of the Basque Country), filled the bilingualism questionaire for French-Basque.

10.3. RESULTS

An omnibus ANOVA analysis was carried out involving the factor Condition (5 levels, one per condition). The omnibus ANOVA analysis was carried out for each variable and only when the main effect of Condition reached a significant value, pairwise comparisons between conditions were computed.

10.3.1. Whole sentence processing and comprehension task

The omnibus ANOVA for the variable *Total Reading Times* for the five *Conditions* showed a significant effect ($F(4,92) = 16.6$, $p < 0.001$). The pairwise comparisons split the conditions into two groups, the singular conditions and the plural conditions (see Fig. 10.6A). The comparison between S_{SING}-V-O_{SING} (3210 ms) and O_{SING}-V-S_{SING} (3163 ms) both singular conditions was not significant ($t(23) = 0.6$, $p > 0.1$). The pairwise comparison between both plural conditions (S_{PL}-V-O_{PL} = 3643ms; O_{PL}-V-S_{PL} = 3669 ms) did not show any statistical difference ($t(23) = -0.3$, $p < 0.7$). On the other hand, the direct comparisons of the singular conditions with the plural counterparts were very significant (S_{SING}-V-O_{SING} vs. S_{PL}-V-O_{PL} $t(23) = -5.9$, $p < 0.001$; O_{SING}-V-S_{SING} vs. O_{PL}-V-S_{PL} $t(23) = -5.8$, $p < 0.001$). Besides, the mean reading time for the fully ambiguous condition (AMB-V-AMB = 3661 ms) behaved like the plural sentences (S_{SING}-V-O_{SING} vs. AMB-V-AMB $t(23) = -4.7$, $p < 0.001$; O_{SING}-V-S_{SING} vs. AMB-V-AMB $t(23) = -4.1$, $p < 0.001$; S_{PL}-V-O_{PL} vs. AMB-V-AMB $t(23) = 0.2$, $p < 0.8$; O_{PL}-V-S_{PL} vs. AMB-V-AMB $t(23) = -0.1$, $p < 0.9$).

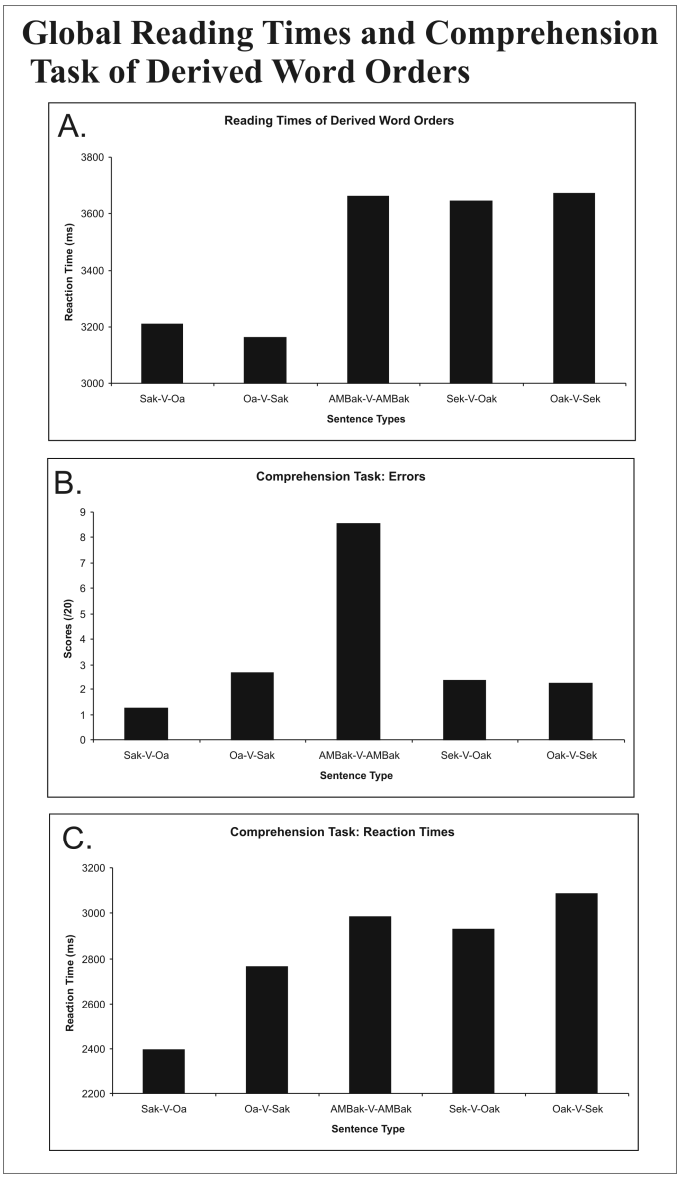

Fig. 10.6*: A: Mean reading times of the five verb-medial derived conditions. B: Means of the performance of the Comprehension Task. C: Error rates (out of 20) of the Comprehension Task of the five experimental conditions.*

The omnibus ANOVAS carried out for the *Comprehension Task* (reaction time and amount of errors) showed main effects of *Condition* in both cases (mean reaction time comprehension task, $F(4,92) = 15.2$, $p < 0.001$; Amount of errors in

the CT $F(4,92) = 43.6$, $p < 0.001$). Fig. 10.6B depicts the mean reaction time in the comprehension task for the different conditions. The questions about S_{SING}-V-O_{SING} condition were faster than any other experimental condition (S_{SING}-V-O_{SING} vs. O_{SING}-V-S_{SING} $t(23) = -4.1$, $p < 0.001$; S_{SING}-V-O_{SING} vs. S_{PL}-V-O_{PL} $t(23) = -6.3$, $p < 0.001$; S_{SING}-V-O_{SING} vs. O_{PL}-V-S_{PL} $t(23) = -6.7$, $p < 0.001$; S_{SING}-V-O_{SING} vs. AMB-V-AMB $t(23) = -6.7$, $p < 0.001$). Comparing O_{SING}-V-S_{SING} to O_{PL}-V-S_{PL} the difference was also significant ($t(23) = -3.0$, $p < 0.007$). The difference between O_{SING}-V-S_{SING} and S_{PL} -V- O_{PL} was not significant ($t(23) = -1.9$, $p > 0.07$). Finally, comparing O_{SING}-V-S_{SING} to the ambiguous condition (AMB-V-AMB) the difference in reading and answering of the comprehension task did not reach a significant level (considering the Bonferroni corrected p-value) ($t(23) = -2.8$, $p > 0.009$).

For the analysis of the amount of errors in the comprehension task, SVO and OVS conditions showed a statistical difference when the constituents of the sentences were singular (S_{SING}-V-O_{SING} vs. O_{SING}-V-S_{SING} $t(23) = 3.1$, $p = 0.005$). Nevertheless, when the constituents were plural the error rate did not show any significance (S_{PL}-V-O_{PL} vs O_{PL}-V-S_{PL} $t(23) = 0.2$, $p > 0.8$). The rest of pairwise comparisons between SVO and OVS conditions were not significant. In the fully ambiguous condition, recall that the error rate of the comprehension task corresponded to the SVO or OVS interpretation of the ambiguous chain. The results of the comprehension task of fully ambiguous condition did not show any preference for SVO nor OVS interpretation (AMB-V-AMB = 42.5%, see Fig. 10.6C).

Summarizing, we have seen that Basque speakers processed sentences containing singular constituents faster, but no difference was found between the SVO and OVS word orders. On the other hand, the comprehension task showed that ambiguous chains did not showed any preference favoring SVO or OVS interpretation. In order to shed further light on the differences we observed across these conditions, we carried out in the next section a word by word comparison for each constituent.

10.3.2. Word-by-word analysis

At sentence initial position (DP1), the omnibus ANOVA was significant ($F(4,92) = 6.2$, $p = 0.001$). Comparing word by word S_{SING}-V-O_{SING} and O_{SING}-V-S_{SING} conditions, the difference did not survive the Bonferroni correction at the first word position ($t(23) = 2.4$, $p < 0.026$), although a trend is observed that points out that the singular absolutive O_{SING} was processed faster than the ambiguously marked S_{SING} (see Fig. 10.7A). Notice that this difference has also been reported in previous studies (see Fig. 10.3, and Erdocia et al. 2009), which could be considered a replication of the previous finding although the marginal level of significance. When comparing the O_{SING} with the first constituents of the other

conditions the difference was significant (O_{SING} vs. S_{PL} $t(23) = -4.0$, $p < 0.001$; O_{SING} vs. O_{PL} $t(23) = -3.7$, $p < 0.001$; O_{SING} vs. AMB $t(23) = -2.8$, $p = 0.009$)[6].

At sentence second position (V), the omnibus ANOVA did not show any significance ($F(4,92) = 0.5$, $p < 0.7$). Thus, no furthered pairwise comparisons were carried out.

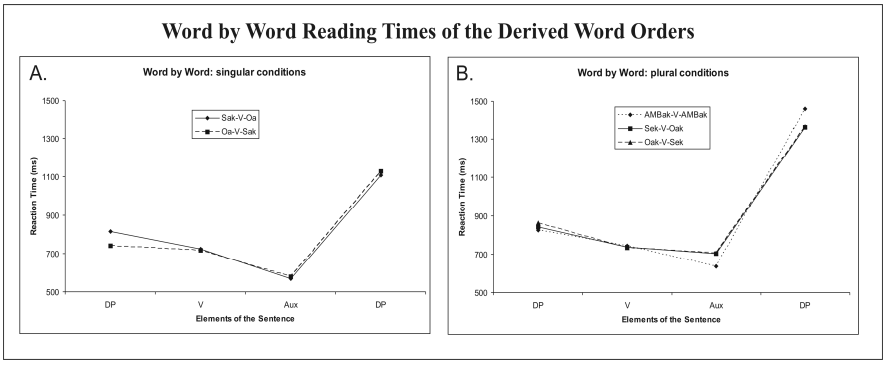

Fig. 10.7: *Word by word comparison of SVO and OVS. A: Mean reading times of every words in singular conditions. B: Means of every word in plural and fully ambiguous conditions.*

At the auxiliary position, the omnibus ANOVA showed a significant effect ($F(4,92) = 18.8$, $p < 0.001$). As Fig. 10.7 shows, the fastest auxiliary was '*du*' in S_{SING}-V-O_{SING} (568 ms) and O_{SING}-V-S_{SING} (578 ms), then the auxiliary of the fully ambiguous condition '*ditu*' (635 ms), and finally the auxiliary with plural agreement with subject and object '*dituzte*' in S_{PL} -V- O_{PL} (701 ms) and O_{PL} -V-S_{PL} (708 ms). The pairwise comparisons revealed that the differences between the conditions with the same auxiliary were not significant (S_{SING}-V-O_{SING} vs. O_{SING}-V-S_{SING} $t(23) = 0.6$, $p > 0.5$; S_{PL}-V-O_{PL} vs. O_{PL}-V-S_{PL} $t(23) = 0.3$, $p > 0,8$). However, the difference between singular (S_{SING}-V-O_{SING} and O_{SING}-V-S_{SING}) auxiliary and AMB auxiliary was significant (*du* vs. *ditu* $t(23) = -3.3$, $p < 0.003$), the difference between singular auxiliary and plural (S_{PL}-V-O_{PL} and O_{PL}-V-S_{PL}) auxiliary was significant (*du* vs. *dituzte* $t(23) = -5.5$ $P < 0.001$) and the difference between AMB and plural conditions (S_{PL}-V-O_{PL} and O_{PL}-V-S_{PL}) was also significant (*ditu* vs. *dituzte* $t(23) = -3.6$, $p < 0.002$). It could be argued that these differences between auxiliaries could be due to their differences in letter length or to their differences in morphological complexity.

Finally, at sentence final position (DP2), the omnibus ANOVA analysis was significant ($F(4,92) = 10.2$, $p < 0.001$). The pairwise comparison revealed that second DP of S_{SING}-V-O_{SING} and O_{SING}-V-S_{SING} conditions had similar mean reading times ($t(23) = -0.4$, $p > 0.7$), as well as for the S_{PL} -V- O_{PL} and O_{PL} -V- S_{PL}

[6] The comparison between the O_{SING} and the AMB did not survive the Bonferroni correction procedure.

conditions ($t(23) = 0.1$, $p > 0.95$). Fully ambiguous condition and conditions with plural constituents also showed similar reading times (S_{PL}-V-O_{PL} vs. AMB-V-AMB $t(23) = 1.3$, $p > 0.2$; O_{PL}-V-S_{PL} vs. AMB-V-AMB $t(23) = 1.1$, $p > 0.3$). The pairswise comparisons of S_{SING}-V-O_{SING} condition with S_{PL}-V-O_{PL} condition ($t(23) = -4.6$, $p < 0.001$), O_{PL}-V-S_{PL} condition ($t(23) = -4.4$, $p < 0.001$), and AMB-V-AMB condition ($t(23) = -4.7$, $p < 0.001$) were significant. Similarly, the comparisons of O_{SING}-V-S_{SING} with S_{PL}-V-O_{PL} condition ($t(23) = -3.6$, $p < 0.002$), O_{PL} -V- S_{PL} condition ($t(23) = -3.7$, $p < 0.001$), and AMB-V-AMB condition ($t(23) = -3.6$, $p < 0.001$) were significant.

10.4. DISCUSSION

10.4.1. Whole sentence processing

The present study revealed that verb-medial SVO and OVS sentences in Basque, do not display a processing asymmetry; in particular, we did not find an advantage for the Subject initial order (SVO) as compared to the Object initial order (OVS). These results converge with claims made in Basque Linguistics that all word orders other than SOV are derived via syntactic displacement. However, studies of word order processing in German, another SOV language, report an advantage of Subject initial (SVO) versus Object initial (OVS) orders. In German, where canonical order is SOV, and V2 applies in main sentences, Bader and Meng (1999) showed that participants were consistently garden-pathed in OSV temporally ambiguous structures and favored the SOV interpretation. Using ERPs, Rösler et al. (1998) showed that sentences whose word order deviates from the SOV canonical structure were more difficult to process (see also Bornkessel et al., 2002, Schlesewsky et al., 2003 among others). In verb medial sentences, ERPs differentiated between SVO and OVS structures in German (Matzke et al., 2002). In Basque, behavioral and ERP measurements differ significantly for canonical SOV and derived OSV word orders (Erdocia et al., 2009): reading times showed longer reading times and syntactic reanalysis effects for OSV structures (see also Fig. 10.2), and ERPs showed negativities in both DPs and P600 at verb position suggesting that derived OSV word order required further computational resources (Erdocia et al., 2009; Fig. 10.3 in this chapter). Thus, while German and Basque reveal a similar processing advantage for SOV versus OSV, an advantage that has also been reported for Japanese, also SOV[7], processing of verb-medial word

[7] In Japanese whose structure is always verb-final some behavioral studies found differences between SOV and OSV structures (Miyamoto & Takahashi, 2002; Mazuka et al., 2002), but some others did not find any difference (Yamashita

orders yields diverging results: in German, a Subject-initial advantage has been reported, whereas in Basque both Subject-initial and Object-initial orders reveal equivalent measures of processing complexity. Arguably, this difference might be due to the fact that German is a V2 language, whereas Basque is not. In V2 languages, main sentences have the inflected auxiliary or verb always in second position in the sentence, and the most frequent constituent preceding the inflected verb is the Subject; this entails that SVO is in fact a rather frequent word order for a V2, main sentence in German, whereas this is not the case for Basque. This explanation is consistent with the results from German, showing a preference of SOV versus OVS, and then a preference for SVO versus OVS, whereas Basque displays a preference for SOV versus OVS, but no preference in the case of SVO and OVS.

The differences we found between singular and plural conditions were arguably due to conceptual and/or morphological complexity, but clearly not to syntactic structure. From a syntactic standpoint, sentences that differ only in having singular or plural subjects are identical, and therefore so were the S_{SING}-V-O_{SING} and S_{PL}-V-O_{PL} sentences in our Basque materials. Hence, the reading time differences must be due to something other than the processing of syntactic structure. Two candidates for the asymmetry encountered are the conceptual representation, which is more complex for plural entities and morphology, given that plural morphology is marked both phonologically and semantically. In Basque, like in many languages, plural forms carry extra morphemes. These morphemes increase the complexity of morphological processing and the length of the words. Finally, fillers could also have had an impact in the differences between singular and plural conditions. As it is shown in examples (13-14) all DP-constituents of the filler sentences were singular.

4.2. Word by word reading times

The word by word analysis we carried out informed about the sources of the differences between conditions. In first word position, the reading times showed that the fastest processed element was singular Object (O_{SING}) (see also Erdocia et al., 2009; and Fig. 10.3 in this chapter). This result diverges sharply from what is usually obtained with the same self-paced reading method in other languages such as Dutch (Kaan, 1997) and Russian (Sekerina, 1997). In these languages Subjects are always processed faster than Objects. This divergent result could be due to one language-particular property: Basque is an ergative language, whereby Objects and intransitive Subjects belong in the same morphological class, whereas transitive Subjects carry an extra morpheme (*-k*), and constitute a class on their own. This extra morpheme on transitive Subjects could be the cause of the increase in

---

1997; Tamaoka et al., 2003). Using ERPs, the differences between canonical SOV and derived OSV arose in Japanese (Hagiwara et al., 2007; Wolff et al., 2008).

processing time for S_{SING}/S_{PL} and the ambiguous condition that could be interpreted either as a singular Subject or a plural Object. Moreover, given ergativity, a sentence initial O_{SING} can be hypothesized to be an intransitive Subject, thus participants can postulate a simple intransitive sentence-structure, and which would in turn yield faster processing.

At auxiliary position (third position), the morphologically most complex auxiliary (*d-it-u-zte* "3rd person/object-plural-root-plural/3rd person/subject") required longest reading time, and the morphologically simplest auxiliary (*d-u* "3rd person/object-root-3rd person/subject") required the shortest reading time regardless of the word order (SVO or OVS) of the condition (Fig. 10.7). The auxiliary of the ambiguous condition fell in-between the most and the less complex auxiliaries. These differences between inflected auxiliaries could be also accounted for in terms of word length, which corresponds with morphological complexity: the shortest auxiliaries were processed faster, and the longest auxiliaries were processed slowest. Whatever the factors responsible for the differences between auxiliaries, syntactic structure is not one of them.

At sentence-final position reading times increased similarly in all five conditions. The pro-drop property of Basque (see Sect. 10.1.3 in this chapter) could be one factor behind such increase in reading times: since both Subjects and Objects can be omitted in Basque, once participants read the auxiliary verb of a sentence, they could assume that the sentence was finished, and that one of the "missing" arguments had been pro-dropped. Leaving aside this overall increase of reading times, the differences that result among conditions at sentence final position, generally replicate the relative differences found at auxiliary position, and they are arguably due to the same reasons that provoked reading time differences at auxiliary position: plural auxiliaries take longer to read than singular ones. However, the longer reading time obtained at second DP position of the AMB-V-AMB condition cannot be the result of the auxiliary taking longer to read; on the contrary, the auxiliary in the AMB condition (*ditu* plural-O/singular-S) was simpler and faster than the auxiliary in S_{PL}-V-O_{PL} condition (*dituzte* plural-O/plural-S). The longer reading time that obtains at sentence final position of ambiguous sentences is arguably due to the participants' realizing at that point that they were reading a fully ambiguous sentence. This interpretation is reinforced by the participants' performance in the comprehension task discussed below. Unfortunately no significant interactions result between condition and word when comparing ambiguous sentences and S_{PL}-V-O_{PL} sentences at auxiliary and second DP positions (S_{PL}-V-O_{PL} vs. AMB-V-AMB, $F(1,23) = 3.6$, $p = 0.071$).

4.3. Comprehension Task

The mean reaction time in the comprehension task revealed that the singular Subject-first condition (S_{SING}-V-O_{SING}) required shorter reading times. The task was performed faster in singular Subject-first condition than in any other condition. Unfortunately, this result could have been biased by the SVO structure

of the comprehension questions. The SVO order of these questions could have facilitated the task for SVO experimental sentences. This explanation, however, seems doubtful for it does not carry to the plural subject-first condition (S_{PL}-V-O_{PL}), which should also show a facilitation effect but does not, despite the fact that the questions of the comprehension task also had the same SVO word order.

Regarding errors in the comprehension task, only the comparison between S_{PL}-V-O_{SING} and O_{SING}-V-S_{PL} showed a significant advantage for the subject-initial condition. Interestingly, the responses to the comprehension task in the fully ambiguous condition did not show any preference for SVO or OVS interpretation. Participants performing the comprehension task were probably aware of the ambiguous character of these sentences. In a study about Japanese (Tamaoka et al., 2003) participants did not show any reaction time difference reading SOV versus OSV structures, but in a comprehension task OSV sentences were harder to judge as correct than canonical sentences. It is possible that speakers realized that they confronted ambiguous sentences while performing the comprehension task, and not before, while reading the sentences.

10.5. CONCLUSION

The aim of the present study was to measure the syntactic complexity and the ambiguity resolution processes that word order variation generates in two syntactically derived structures: verb-medial SVO and OVS. By using *self-paced reading* and *comprehension task*, our data showed that both verb-medial word orders were equally processed by native Basque speakers. This result suggests that both derived word orders required similar computational resources, with no advantage for the subject-before-object sequence. On the other hand, our data did not indicate any ambiguity resolution process reading temporally ambiguous sentences and fully ambiguous chains.

ACKNOWLEDGEMENTS

This research has been supported by grants from the Spanish Ministry of Education and Science (CSD2007-00012), the Spanish Ministry of Science and Innovation (FFI2009-09695), and the Basque Council for Education, Universities and Research (IT414-10) to Itziar Laka and Kepa Erdocia, and the Spanish Government (SEJ2005-06067/PSIC) to Antoni Rodriguez-Fornells.

REFERENCES

Artiagoitia, X. 1995. Verbal Projections in Basque and Minimal Structure. Supplements of Anuario del Seminario de Filología Vasca <<Julio de Urquijo>>, XXXVI.

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

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

Boskovic, Z; Takahashi, D. 1998. Scrambling and Last Resort. Linguistic Inquiry, 29, 347-366.

Chomsky, N. 1957. Syntactic Structures, Mouton, The Hague.

Chomsky, N. 1981. Lectures on government and binding. Foris: Dordrecht.

De Rijk, R. 1969. Is Baque an S.O.V. language? Fontes Linguage 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.

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

Erdocia, K.; Laka, I.; Mestres-Missé, A.; Rodriguez-Fornells, A. 2009. Syntactic complexity and ambiguity resolution in a free word order language: behavioural and electrophysiological evidences from Basque. Brain and Language, vol. 109 (1), 1-17. :

Etxepare, R.; and Ortiz de Urbina, J. 2003. Focalization. In A Grammar of Basque, José Ignacio Hualde and Jon Ortiz de Urbina (eds.), 459-516. Mouton de Gruyter: Berlin.

Fernandez, B. 1998. Egiturazko kasuaren erkaketa euskaraz. PhD dissertation. EHUko argitalpen zerbitzua: Bilbo.

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.

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. Ph.D. Thesis: University of Groningen (Groningen Dissertations in Linguistics 20).

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.

Knoeferle, P.; Habets, B.; Crocker, M.; Münte, T. 2008. Visual Scenes Trigger Immediate Syntactic Reanalysis: Evidence from ERPs during Situated Spoken Comprehension. Cerebral Cortex, 18, 789-795.

Laka, I. 1990. Negation in Syntax: On the Nature of Functional Categories and Projections. MIT Dissertation. MITWPL: Cambridge, MA.

Laka, I. 1993. "The Structure of Inflection: A Case study in X^0 Syntax" In Generative Studies in Basque Linguistics [Current Issues in Linguistic Theory 105], J.I. Hualde & J. Ortiz de Urbina (eds.) 21-70, Amsterdam/Philadephia: John Benjamins.

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

Laka, I. 2006. "On the nature of case in Basque: structural or inherent?" en Hans Broekhuis, Norbert Corver, Jan Koster, Riny Huybregts and Ursula Kleinhenz & Jan Koster (eds.), Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk, Berlin/New York, Mouton de Gruyter, pp. 374-382, ISBN 3-11-018850-3.

Levin, B. 1983. *On the Nature of Ergativity*, PhD Dissertation, MIT, Cambridge, Masschusetts, USA.

McCarthy, G.; and Wood., C. 1985. Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. Electroencephalography and clinical Neurophysiology; 62, 203-208.

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

Mazuka, R.; Itoh, K.; 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.; 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.

Ortiz de Urbina, J. 1989. Parameters in the Grammar of Basque. Foris: Dordrecht.

Ortiz de Urbina, J. 2003. Word Order. In José Ignacio Hualde and Jon Ortiz de Urbina (eds.). A Grammar of Basque. 2003. Mouton de Gruyter: Berlin. 448-459.

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.

Rebuschi, G. 1987. Sujeto, ergatividd y (no-) configuracionalidad; lugar del euskara en la tipología generativa. In Euskal Morfosintaxia eta Fonologia: Eztabaida gaiak. Pello Salaburu (ed.), EHU, Bilbo, 9-56

Robins, R. H. 1967. A Short History of Linguistics. Longman: London.

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.

Sekerina, I. 1997. The syntax and processing of Russian scrambled constructions. Unpublished Ph.D Dissertation. City University of New York: NY.

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.

Wolff, S.; Schlesewsky, M; Hirotani, M; Bornkessel-Schlesewsky, I. 2008. The Neural Mechanisms of Word Order Processing revisited: Electrophysiological Evidence from Japanese. Brain and Language, 107, 133-157.

Yamashita, H. 1997. The Effects of Word Order and Case Marking Information on the Processing of Japanese. Journal of Psycholinguistic Research, 26, 163-188.