

Asymmetry in child comprehension and production of Basque subject and object relative clauses

M. Juncal Gutierrez-Mangado and Maria-José Ezeizabarrena
University of the Basque Country¹

1. Introduction

Research on the acquisition of relative clauses by Typically Developing (TD) as well as by SLI children in a variety of languages has revealed that SRs (e.g. *the girl that is painting the lady*) are produced and comprehended earlier and more easily than ORs (e.g. *the girl that the lady is painting*): English (Brown, 1972; Hamburger and Crain, 1982; Grodzinsky, 1989; de Villiers et al., 1994; van der Lely, 1996; McKee et al., 1998), Swedish (Håkansson & Hansson, 2000), Greek (Stavrakaki, 2001), Hebrew (Friedmann & Novogrodsky, 2004; Friedmann et al., 2009), Italian (Arosio et al., 2005), Korean (Cho, 1999), Quechuan (Courtney, 2006), among many others.

It has been observed that the difficulty in producing and comprehending ORs is modulated by differences in the feature make up of the NPs involved in ORs such (i) a mismatch between the case marking (or syntactic function on the head and the gap (Sauerland & Gibson 1998), (ii) differences in the nominal restriction of the NP inside the RC and the head of the RC (Friedmann et al. 2009) and (iii) number and gender mismatch between the NP inside the RC and the head of the RC (Adani et al. 2010, among others). ORs where there is a feature mismatch in terms of nominal restriction, case, number and gender are processed with greater accuracy than those where there is no such mismatch. The animacy configuration of the nouns involved in a RC also affects the ease of processing of ORs. In this respect, Mak et al. (2002), Traxler et al. (2002) among others, reported that ORs with inanimate heads were processed faster than those with animate heads in a reaction times and eye-tracking experiments respectively. This finding has also been replicated in L1 (Bever 1970, Correa 1995, Ozeki & Shirai 2007) as well as in L2 acquisition (Ozeki & Shirai 2007, Kanno 2007).

However, it is noteworthy that most of this work has been based on (i) SVO nominative–accusative languages (Brown 1972, Hamburger & Crain, 1982, McKee, McDaniel & Snedeker, 1998, Friedmann & Novogrodsky, 2004, Friedmann, Belletti & Rizzi, 2009, among others) and (ii) in the acquisition of languages with postnominal

¹ Acknowledgements: The authors are very grateful to the adults and children that participated in the study, as well as to Ane Berro, Aranzadi Ikastola in Bergara, Axular Ikastola in San Sebastian, and Karmelo Ikastola in Bilbao. Institutional and financial support by the University of the Basque Country (GIU09-39 and 00103.130-13578/2001), the Ministry of Science and Technology (CSD2007-00012 and BFF2000-0101), the Ministry of Science and Innovation (FFI2009-13956-C02-01) and the Basque Government (PI2009-22 and IT-210-07; UFI11/06) made this research possible.

relatives. Results from the acquisition of languages with prenominal RCs have yielded mixed results. Some studies have reported a SR preference (Lee (1991) and Cho (1999) for Korean and Chang (1984), Lee (1992), Cheng (1995), Su (2004) and Hsu et al. (2009) for Mandarin Chinese), while others have reported an OR advantage, (Hakuta (1981) for Japanese and Yip and Mathews (2007) for English-Korean bilingual children) or no difference in emergence between SRs and ORs (Ozeki & Shirai (2005, 2007) for Japanese).

The present paper focuses on the production of RCs in Basque, a highly inflected, ergative SOV language with prenominal relatives and compares these results with previous comprehension data in the same language.

2. Basque Relative Clauses

Basque is a highly inflected, ergative, SOV language where the inflected verb form (the auxiliary in most cases) agrees (in person, number and case) with the subject, direct object and indirect object, as can be seen in (1-2) below (Altube 1929, de Rijk 1969, Salaburu 1987, Ortiz de Urbina 1989, among many others):

In intransitive sentences, the subject bears the zero marked absolutive (A) case as in (1), the same as the one borne by the DO *egunkaria-Ø* ‘the newspaper’ in (2a-d).

- (1) *Gizon-a- Ø etorri- Ø da*
 child-the-A come-PF aux-3sA
 ‘the man has come’

The subject of the transitive sentence in (2) bears the ergative (E) case marker *-k* (2a-d). The indirect object *gizonari* ‘to the man’, bears the definite suffix *-a* and the dative (D) case marker *-ri*. The direct object, *egunkaria* ‘newspaper’, bears the definite suffix *-a* and the absolutive (A) case marker \emptyset . Arguments may be placed in almost all logical orders in this free word order language (2a-d). Moreover, besides allowing the dropping of the subject, as many highly inflected languages, Basque also allows the dropping of direct and indirect objects (2e):

(2)

- a. *Zu-k gizon-a-ri egunkaria-a-Ø eman-Ø d-i-o-zu*
 you-E man-the-D newspaper-the-A give-PF 3sA-aux-3sD-2sE
 ‘You have given the man the newspaper.’
- b. *Gizon-a-ri egunkaria-a-Ø zu-k eman-Ø d-i-o-zu*
- c. *egunkaria-a-Ø gizon-a-ri zu-k eman-Ø d-i-o-zu*
- d. *zu-k egunkaria-a-Ø gizon-a-ri eman-Ø d-i-o-zu*

- e. *eman-Ø d-i-o-zu*
 give-PF 3sA-aux-3sD-2sE
 ‘You have given it to him/her’

Typically, RCs in Basque lack a *wh*-element heading the relative clause. The RC is marked by adding the subordinating suffix *-en* to the auxiliary of the relative clause and it precedes the noun it modifies, that is, restrictive SRs (3) and ORs (4) in Basque are prenominal (Oyharçabal 1987, Ortiz de Urbina 1990, Artiagoitia 1992):

- (3) *Hau da [e₁ amon-a-Ø muxuka-tzen du-en] neska-a-Ø₁*
 this is [__E grandmother-the-A kiss-IPF aux-REL girl-the-A
 ‘This is the girl who is kissing the grandmother.’

- (4) *Hau da [amon-a-k e₁ muxukatzen du-en] neska-a-Ø₁*
 this is grandmother-the-E __A kiss-IPF aux-REL girl-the-A
 ‘This is the girl who the grandmother is kissing.’

Several characteristics are noteworthy from the examples above: (i) as described above, the head *neska* ‘girl’ follows the RC in brackets, as in Chinese, Korean or Japanese and unlike in English²; (ii) in the SR the subject gap bears ergative case while the antecedent, *neska* ‘girl’, bears absolutive case. That is, in SRs like (3), there is a mismatch between the case in the gap and the case in the head. No such mismatch is found in the OR in (4), where both the object gap and the head bear absolutive case³; (iii) another characteristic of RCs in Basque is that like in other SOV languages, the filler does not precede the gap as in English-like languages. As can be observed above the gap precedes the filler in both SRs and ORs.

A final characteristic of Basque is of relevance for the present purpose: in both production experiments the tense used for eliciting RCs was the present tense. The examples in (3) and (4) above show imperfective aspect in the present tense, which is an appropriate answer for the test items in both production experiments. The imperfective morphological marking may have a habitual as well as a progressive meaning when referring to an event taking place simultaneously to the speech act. This distinction is

² Note also that the head noun can be omitted but is represented by the determiner *-a* which appears attached to the *-en* relativizer in the auxiliary as shown in (i) below, which corresponds to (3) above. We will refer to this structure as a nounless head:

- (i) *Hau da amon-a- Ø muxukatzen du-en-a- Ø*
 this is grandmother-the-A kiss-IPF aux-REL-the-A
 ‘This is the one who is kissing the grandmother’

³ In the case of the SR in (3), there is a mismatch between the case marking of the subject gap, which is ergative, and the case marking in the head noun, which is absolutive. This alone could increase the processing difficulty of SRs when compared to ORs. However, as Carreiras et al. (2010) reported, despite having no such case mismatch in their SR and OR test sentences, the results showed that ORs were processed more easily than SRs.

important for this article, as the aspectual meaning of the inflected forms varies across the experiments. Habituality is the typical meaning of examples like (3) and (4) though in situations like picture descriptions this morphological form may be used to refer to the progressive meaning. There are, however, other grammatical options of inflected V forms for this particular situation, such as (5) and (6). In sentences (3-4) the external argument of the embedded predicate bears ergative case (as can be observed in the overt S of the OR in (4)) and the internal argument of the SR bears absolutive case (3). These typically progressive forms are the ones which we shall call the *ari*⁴ progressive periphrasis. This periphrastic form involves a change in the case assignment of the external argument of the predicate as shown in (5) for an SR and in (6) for an OR:

(5) *Hau da [amona-a-Ø e₁ muxukatzen ari den] neska-a-Ø₁*
 this is [grandmother-the-A kiss-IPF PROG aux-REL girl-the-A
 ‘This is the girl who is kissing the grandmother/This is the girl who the
 grandmother is kissing’

(6) *Hau da [e₁ amona-a-Ø muxukatzen ari den] neska-Ø₁*
 this is [grandmother-the-A kiss-IPF PROG aux-REL girl-the-A
 ‘This is the girl who is kissing the grandmother/This is the girl who the
 grandmother is kissing’

As can be seen from the examples above, when using the ‘ari’ progressive, the RCs are ambiguous between a subject and an object reading because the use of ‘ari’ involves case absorption (for further details see Hualde & Ortiz de Urbina (1987) and Laka (2006)).

3. The study

Three different experiments were carried out in order to investigate the comprehension and the production of RCs in Basque. On the one hand, a comprehension experiment and a production experiment with the same groups of subjects (Experiments 1 and 2 respectively) and a different production experiment with another group of subjects (Experiment 3).

3.1. Participants

Details of the participants can be observed in Table 1:

⁴ Besides *-T(z)EN ARI*, other periphrastic structures have the same effect: *-T(z)EN ibili* ‘be about’ and *-T(z)EN egon* ‘be located’.

	Age 4 (N=13)	Age 5 (N=20)	Age 6 (N=9)	Adults (N=20 ⁵ and N=10)
EXPERIMENT 1 (Comprehension) AND EXPERIMENT 2 (Production)	7 ♀, 6 ♂ M=4.1 SD=3.9m.		5 ♀, 4 ♂ M=6.2 SD=1.6m.	11 ♀, 9 ♂ M=17 SD=5m.
EXPERIMENT 3 (Production)		9 ♀, 11 ♂ M=5.6 SD=3.8 m.		8 ♀, 2 ♂ M=35.7 SD=14 y.

Table 1: Participants in both experiments (M=mean; SD=Standard Deviation; m.= months; y.=years).

3.2 Materials

The same set of materials was used in the Experiment 1 and 2 as shown in Figure 1 below⁶. In the production experiment, the experimenter presented two pictures with two characters each performing the same action: in the first picture one of the characters performed an action on the second character and in the second picture the roles were reversed and the second character became the doer of the action. The predicates were always transitive and the actions were always reversible.



Figure 1: Sample picture in Experiments 1 and 2

The Binary-picture sentence matching task used in the comprehension experiment included 80 sentences: 40 SOV, 20 SR and 20 OR sentences (Friedmann and Novogrodsky, 2004).

(7)

- a. SOV: *Mutiko-a-k pinguinu-a-Ø garbi-tzen du*
boy-the-E penguin-the-A wash-IPF aux
'The boy is washing the penguin'
- b. SR: *[Pinguino-a-Ø garbi-tzen du-en] mutiko-a-Ø*
penguin-the-A wash-IPF aux-REL boy-the-A
'The boy that is washing the penguin'

⁵ The first N refers to the number of participants in Experiments 1 and 2 and the second N to those in Experiment 3.

⁶ We thank Na'ama Friedman for the materials used in Experiments 1 and 2.

- c. OR: [Pinguino-a-k garbitzen du-en] mutiko-a-Ø
 penguin-the-E wash-IPF aux-REL boy-the-A
 ‘The boy that the penguin is washing’

For the first production experiment the same visual materials as the ones used in the comprehension experiment were used. Three conditions were considered: SR (20 sentences), OR (20 sentences) and SOV (40 sentences).

For the second production experiment, on the other hand, no visual material was used. In this case, the experimenter presented two characters performing two actions in which one of the characters is the agent (10 sentence pairs) or the patient (10 sentence pairs), and asked the participant to choose which one would (s)he preferred to be (Friedmann 2010). The verbs were always transitive and two conditions were considered: SR (10 sentences) / OR (10 sentences) and semantic reversibility of the predicates (10 reversible and 10 non reversible).

Experimenter eliciting SRs (8a): *There are two girls. One kisses her mother and the other pushes her mother. What girl would you rather be?* Experimenter eliciting ORs (8b): *There are two girls. One elephant lifts one girl and the elephant wets the other girl. What girl would you rather be?*

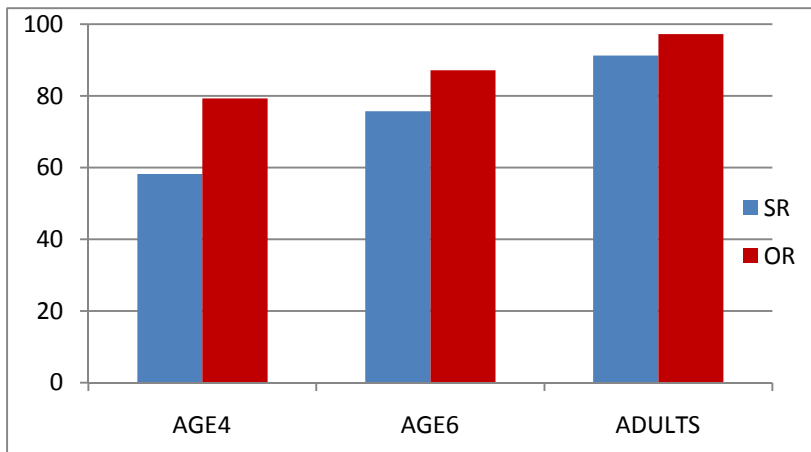
(8)

- a. SR: [amatxo-Ø bultza-tzen du-en] neska-a-Ø
 Mummy-A push-IPF aux-REL girl-the-A
 ‘The girl that pushes her mother’
- b. OR: [elefante-a-k altxa-tzen duen] neska-a-Ø
 elephant-the-E lift-IPF aux-REL girl-the-A
 ‘The girl that the elephant lifts’

4. Results

4.1. Comprehension

The results in Experiment 2 showed that ORs are understood better than SRs in all groups (Gutierrez-Mangado 2011): 4 year-olds: ($z=-2.32$, $p=0.020$), 6 year-olds: ($z= -1.95$, $p=0.051$) and adults ($z= -2.88$, $p=0.004$). This can be observed in Graph 1 below:



Graph 1: Results from the comprehension experiment (Experiment 1)

4.2. Production

The results from both production experiments are presented separately. In Experiment 2 from the total sentences elicited in the adult group, 99.28% (834/840) were RCs, in the 4-year-old group 95,7% (498/520) and in the 6-year-old group 98% (353/360). In Experiment 3 the adults produced a total of 197 RCs out of 199 items (98.9%) and the 89.3 % (352/394) of the sentences produced by the 5-year-old children were RCs. Table 2 below shows the distribution of these RCs:

	Age 4 (N=13)	Age 5 (N=20)	Age 6 (N=9)	Adults (N=20 and N=20)
Experiment 2	95.4% SR (248/260) 96.2% OR (250/260)		98.3% SR (177/180) 97.8% OS (176/180)	99.8% SR (419/420) 98.8% OR (415/420)
Experiment 3		92.4% SR (182/197) 86.3% OR (170/ 197)		100% SR (99/99) 98% OR (98/100)

Table 2: Raw production of RCs across samples

The data from Experiments 2 and 3 are not normally distributed in SRs (Exp 2: $Z=2.90$; $p<.01$ and Exp 3: Kolmogorov-Smirnov $Z=1.56$; $p<.01$) or in ORs (Exp 2: $Z=2.89$; $p<.01$ and Exp 3: $Z=2.43$, $p<.05$). Therefore, the statistical analyses carried out throughout the paper are non-parametric. The Wilcoxon test showed no statistically significant differences between the SRs and ORs produced by any of the different age groups (Exp 2: age 4: $Z=-0.54$, $p>.05$; age 6: $Z=-0.27$, $p>.05$; adults: $Z=-1.13$, $p>.05$; Exp 3: age 5: $Z=1.89$, $p=0.59$; adults: $Z=1.41$, $p=0.16$).

In order to analyse the relative clauses produced, several methodological decisions were made. To start with, from all the RCs produced, those which were ambiguous between a SR and an OR reading were eliminated from the raw data. The ambiguity was caused by

three factors: (i) the lack of explicit overt arguments inside the RC (9), (ii) RCs involving “ari” type structures (10) and (iii) RCs with atypical word order (11). Note that all these RCs were grammatical. Let us examine each of the eliminated RCs.

Absence of an overt S or O argument in the embedded clause makes the coindexation of the head with any of the silent arguments of the embedded predicate possible, resulting in an ambiguous sentence, which can be interpreted as SR or as OR, as shown in (9a) for a RC with an overt head and in (9b) for a RC with a nounless head:

(9)

- a. SR/OR: [_____i _____j *altxa-tzen duen* *neska-a-Ø*_{ij}
 lift-IPF aux-REL girl-the-A
 ‘the girl that lifts something’ / ‘the girl that somebody lifts’

- b. SR/OR: [_____i _____j *altxa-tzen duen*]-a-Ø_{ij}
 lift-IPF aux-REL-the-A
 ‘The one that lifts somebody’/ ‘the one that somebody lifts’

As for RCs with “ari” type periphrases, recall from the discussion above that Basque allows the possibility of forming RCs in the progressive by means of an intransitive inflected auxiliary, like the case of the *-T(z)EN ari* construction or its homologous *-T(z)EN egon* ‘to be ___-ing’ and *-T(z)EN ibili* ‘to be about ___-ing’. Such sentences are ambiguous regardless of the overt (10a) or non-overt (10b) absolutive argument, which can be interpreted as the internal (O) or the external argument (S) of the periphrastic V in the RC.

(10)

- a. SR/OR [(_j) *amama-Ø*_i (_j) *pinta-tzen dago*_{ij}-n]-a-Ø
 grandmother-the-A paint-IPF is-REL-the-A
 ‘the one that is painting the grandma/ that the grandma is painting’

- b. SR/OR [_____i _____j *pinta-tzen dago*_{ij}-n]-a-Ø
 paint-IPF is-REL-the-A
 ‘the one that is painting’ or ‘the one that someone is painting’

Finally, RCs with atypical word order have been eliminated from the general analysis because they can be ambiguously interpreted⁷. In examples such as (11) the DP *mutikua* on the right of the RC could be interpreted in two ways: (i) as the overt object argument of the RC extracted to the right, in which case (11) could be interpreted as a SR or (ii)

⁷ Note that RCs with atypical word order may be regarded as non-standard or even ungrammatical by some speakers.

as a lexical doubling of the head of the RC, in which case the sentence could be interpreted as either a SR or an OR.

- (11) *Garbi-tzen dab-en-a-Ø mutiku-a-Ø*
 wash-IPF aux-REL-the-A boy-the-A
 ‘The one that washes the boy / the boy that washes’
 TARGET : Mutikua garbitzen dabena / garbitzen daben mutikua
 ‘The one that washes the boy / the boy that washes’

Table 3 below shows the percentages of excluded RCs by group:

	Age 4	Age 5	Age 6	Adults exp.2	Adults exp. 3
0 ARG	SR: 3.2% (8/248) OR: 4.8% (12/250)	SR: 2.2% (4/182) OR: 7.64% (13/170)	SR: 1.1% (2/177) OR: 1.7% (3/176)	SR: 0% (0/419) OR: 0% (0/415)	SR: 5% (5/99) OR: 10.2% (10/98)
“ari” TYPE	SR: 1.2% (3/248) OR: 0.8% (2/250)	SR: 5.5% (10/182) OR: 5.88% (10/170)	SR: 10.7% (19/177) OR: 11.3% (20/176)	SR: 57.1% (237/419) OR: 47.2% (198/415)	SR: 0% OR: 0%
Atypical word order	SR: 19.3% (48/248) OR: 14% (35/250)	SR: 7.1% (13/182) OR: 11.1% (19/170)	SR: 22% (39/177) OR: 7.9% (14/176)	SR: 4.7% (20/419) OR: 4.5% (19/415)	SR: 0% OR: 1% (1/98)

Table 3: percentages of excluded RCs by type across the groups

For experiment 2, in order to see whether any of the three excluded types differed within groups the Wilcoxon test was used revealing no statistically significant differences within each group: for *0 Argument* SR vs. OR (Age 4: $z=-1.03$, $p=.302$; Age 6: $z=-5.35$, $p=.593$; Adults: $z=.000$, $p>.005$), for “*ari*” *Type* in SR vs. OR (Age 4: $z=-1.089$, $p=.276$; Age 6: $z=-1.000$, $p=.317$; Adults: $z=-1.734$, $p=.083$) and for *Atypical Word order* differences in SR vs. OR (Age 4: $z=-1.604$, $p=.109$; Age 6: $z=-1.153$, $p=.249$; Adults: $z=-.424$, $p=.672$). For experiment 3, no statistically significant within-group differences in the production of SR or OR was found in any of the three cases: for *0 Argument* (Age 5: $z=-1.89$, $p=.058$; Adults: $z=-0.94$, $p=.344$), for *ari Type* (Age 5: $z=-0.271$, $p=.786$; Adults: $z=-0.0$, $p=1.0$) and for *Atypical Word order* (Age 5: $z=-1.38$, $p=.166$; Adults: $z=-1.0$, $p=.31$).

To carry out the between groups comparison with respect to the production of *0 Argument*, “*ari*” *Type* and *Atypical Word Order* in SRs and ORs, the Kruskal-Wallis test was used in both experiments. The results in Experiment 2 indicated that there were differences between the groups for each type of structure. These differences were between the adult group and 4-year-old group. With respect to *0 Argument*, the 4-year-old group omitted the internal argument of the OR with higher frequency than the adult group ($z=-3691$, $p=.008$). Regarding “*ari*” *type* constructions, the subjects in the adult group used this structure with greater frequency than the 4-year-old group both in SRs ($z=-4.302$, $p<.001$) and in ORs ($z=-3.630$, $p<.001$). Finally with respect to RCs with *Atypical Word Order*, the adults’ production was much lower than the 4-year olds (SR: $z=-2.961$, $p<.01$; OR: $z=-2.843$, $p<.01$). In Experiment 3, the only difference between

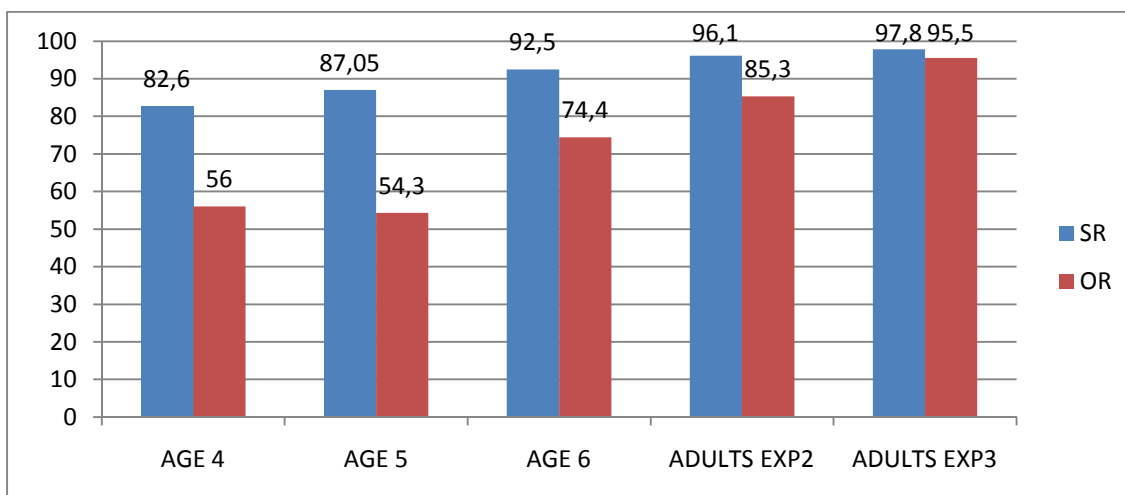
the 5-year-old group and the adults involved the higher frequency of RCs with *Atypical Word Order* of the former in both SR ($z=-2.46, p<.05$) and OR ($z=-2.02, p<.05$).

To sum up, the total percentage of unambiguous RCs included for analysis is shown in Table 4:

	Age 4 (N=13)	Age 5 (N=20)	Age 6 (N=8)	Adults (N=9 and N=10)
Experiment 2	SR: 76.6% (190/248) OR: 80% (200/250)		SR: 84.9% (135/159) OR: 90.3% (141/156)	SR: 86% (154/179) OR: 89.2% (157/176)
Experiment 3		SR: 93.9% (171/182) OR: 88.2% (150/170)		SR: 95.9% (95/99) OR: 86.7% (85/98)

Table 4: Percentage of unambiguous RCs included for analysis across groups

The statistical analyses showed that there were no within or between group differences in the percentage of unambiguous SR and ORs included for analysis in either experiment ($p>.05$ in all samples) which leads us to conclude that the production of ambiguous sentences cannot be interpreted as evidence for an avoidance strategy for either the production of SRs and ORs. Importantly, note that in Experiment 2 one subject had to be eliminated from the 6-year-old group and as many as 11 from the adult group because their RCs were all “ari” type⁸. The unambiguous RCs included for analysis, thus, are those excluding RCs with 0 argument, with “ari” type morphology and those with atypical word order. Graph 2 shows the percentages of target-like SRs and ORs:



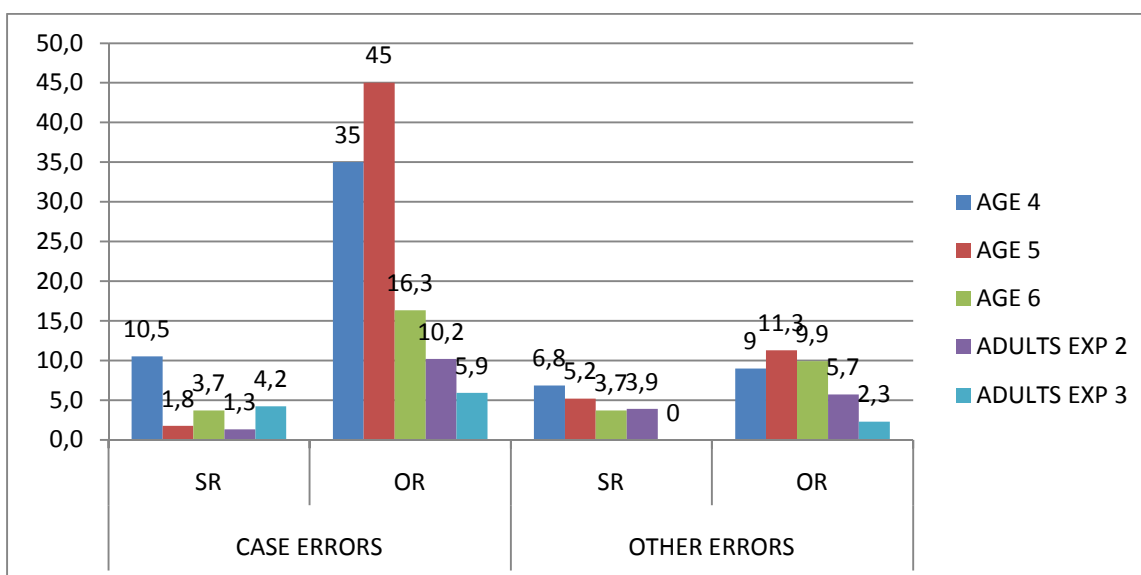
Graph 2: Production of unambiguously target-like SRs and ORs across all groups

As it can be observed in Graph 2, there is an asymmetry between the production of grammatical SRs (range 82-97%) and ORs (range 54-95%), being the percentage of

⁸ This means that in all subsequent analyses the total number of subjects in the Age 6 group is 8 and 9 in the adult group in Experiment 2.

grammatical SRs 15% to 32% higher than ORs in all children's groups and from 2% to 11% higher than ORs in adult productions⁹. The statistical analyses carried out revealed that some of these differences reached significance. In Exp 2 there were within group differences in the production of grammatical SRs and ORs in the 4-year-old group ($z=-2.55$, $p<.05$) and in the adult group ($z=-2.37$, $p<.05$) and in Exp 3 in the 5-year-old group ($z=-2.16$, $p<.05$). Regarding between group comparisons, in Exp 2 the only difference found was between the 4-year-old group and the adult group ($z=-3.62$, $p<.01$) in the production of SRs ($z=-2.73$, $p<.01$) and ORs ($z=-2.77$, $p<.01$) and between the 5-year-old group and the adults only in ORs ($z=-4.04$, $p<.01$) in Exp 3.

In what follows we will present the main types of errors produced in the production of RCs in order to account for these differences. To start with, as the results shown in Graph 3 below show, the major source of error was the production of the incorrect case marker, namely the omission of the ergative case marker *-k* in ORs:



Graph 3: Error distribution across samples

Let us now look at these errors in more detail beginning with the most common error:

(i) Case errors: the overt internal argument of the RC shows incorrect case marking. As a consequence of this, what should be an OR, where the internal argument of the RC must take ergative case, shows up as an SR, where the internal argument bears absolutive case, as shown in (12a) and vice versa as in (12b):

⁹ The adults in Exp2 produced 15 instances of ORs featuring the anti-passive construction (one of the subjects producing 7), as the one shown in (i) below. No such examples were found in the children groups.

(i) *errege-a-k orraztu-a izan d-en-a-∅*
king-the-E comb-PF be aux-REL-the-A
‘The one that is combed by the king’

SR instead of OR:

- (12) a. *Mutiko-a-∅ sika-tzen du-en-a-∅*
boy-the-A dry-IPF aux-REL-the-A
'The one that is drying the boy'
TARGET: *Mutikoak sikatzen duena*
'The one that the boy is drying'

OR instead of a SR:

- b. *Amatxo-k sika-tzen dab-en-a-∅*¹⁰
Mummy-E dry-IPF aux-REL-the-A
'The one that mummy is drying'
TARGET: *Amatxo sikatzen dabena*
'The one that is drying mummy'

(ii) The second type of error involves RCs with incorrect verbal agreement. Most of these errors involve the use of the wrong verbal agreement affix, usually involving the production or absence of the dative marker in the auxiliary as shown in (13a) or cases in which the verb agrees with the first person singular like in *I would rather be ...* as in (13b):

(13)

- a. *Erregi-a-i hola-n itten dab-en-a-∅*
king-the-D this-like do-IPF aux-REL-the-A
'The one that is doing like this to the king'
TARGET: *Erregiai holan itten dotzana*
'The one that is doing like this to the king'
- b. *Helatu-a-∅ jat-en dut-en-a-∅*
ice-cream-the-A eat-IPF aux1s-REL-the-A
'The one that I eat the ice-cream'
TARGET: *helatua jaten duena*
'The one that eats the ice-cream'

(iii) The third type of error involves the use of resumptive pronouns. A few of the sentences featured the use of a resumptive pronoun. Example (14a) shows a resumptive NP in an OR and (14b) an OR with a resumptive pronoun:

¹⁰ Morphological variability can be observed across some of the examples produced by the participants. These differences, corresponding to local varieties of Basque, are irrelevant to the morphosyntactic properties of RCs.

(14)

- a. *Behixa-a-k jirafa-i txupa-tzen dotza-n-a-Ø*
Cow-the-E giraffe-the-D lick-IPF aux-REL-the-A
‘The one that the cow is licking to the giraffe’
TARGET: *Jirafai txuputzen dotzana* or *Jirafa txuputzen dabena*
‘The one (cow) that is licking the giraffe’
- b. *Amatxo-k bera-i lehor-tzen dotza-n-a-Ø*
Mummy-E her-D dry-IPF aux-REL-the-A
‘The one that the mother is drying to her’
TARGET: *Amatxok lehortzen dotzana* or *amatxok lehortzen dabena*
‘The one that mummy is drying’

Finally, some RCs showed a change in the thematic roles (transitivity) of the predicate, as illustrated in (15):

- (15) *argazkixa-a-Ø atera-tzen dab-en-a-Ø*
photograph-the-A take-IPF aux-REL-the-A
‘The one taking the photograph’
TARGET: *erizaina filmatzen duen neska*
‘The girl that is filming the nurse’

It must be noted that in Experiment 2 this type of error occurred with the predicate ‘film’ mostly. A further set of RCs were considered as errors as they involved a theta role reversal:

- (16) *umi-a-k marraz-ten dab-en-a-Ø*
child-the-E draw-IPF aux-REL-the-A
‘The one the child is drawing’
TARGET: *amak marrazten duen neska*
‘The girl the mother is drawing’

Table 7 below shows the percentage of these errors in the production experiments:

	Age 4	Age 5	Age 6	Adults exp.2	Adults exp. 3
Case	SR: 10.5% (20/190) OR: 35% (70/200)	SR: 1.76% (3/171) OR: 45% (30/150)	SR: 3.7% (5/135) OR: 16.3% (23/141)	SR: 1.2% (2/154) OR: 10.1% (16/127)	SR: 4.2% (4/95) OR: 5.9 (5/85)
Agreement error	SR: 1.5% (3/190) OR: 1.5% (3/200)	SR: 4.7% (8/171) OR: 4.7% (7/150)	SR: 0% OR: 2.8% (4/141)	SR: 1.2% (2/154) OR: 0.63% (1/127)	SR: 0% OR: 0%
Resumptives	SR: 2.6% (5/190) OR: 6.5% (13/200)	SR: 0 OR: 0.6% (1/150)	SR: 1.4% (2/135) OR: 3.5% (5/141)	SR: 0.6% (1/154) OR: 0.6% (1/127)	SR: 0% OR: 1.17% (1/85)
Transitivity change	SR: 3.1% (4/190) OR: 8.9% (1/200)	SR: 0.6% (1/171) OR: 3.3% (5/150)	SR: 1.4% (2/135) OR: 1.4% (2/141)	SR: 0.6% (1/154) OR: 2.5% (4/127)	SR: 0% OR: 1.17% (1/85)
Head or theta role reversal	SR: 0.5% (1/190) OR: 0.5% (1/200)	SR: 0% OR: 2.7% (4/150)	SR: 0.7% (1/135) OR: 2.1% (3/141)	SR: 1.2% (2/154) OR: 1.9% (3/127)	SR: 0% OR: 0%

Table 7: Error types and frequencies across the groups

Most kinds of errors, like S-V agreement, transitivity change and theta role reversal are scarce as most of them show values below 5% in the majority of samples. In contrast, case errors are quite consistent (range 45% to 5% across samples) in ORs and much less frequent in SR (range 10% to 1% across samples).

For Experiment 2, in order to find out if there were any within-group differences in the production of these errors, we carried out the Wilcoxon test. The results showed that in each of the groups the production of errors in SRs and ORs only differed in Case errors, where the subjects tested failed to produce the correct case for the internal argument in ORs with higher frequency than in SRs (Age 4: $z=-2.707$, $p<.05$; Age 6: $z=-2.207$, $p<.05$; Adults: $z=-2.213$, $p<.05$). In order to determine whether the types of errors differed between groups we carried out the Kruskal-Wallis test. The results revealed that the groups were different. Once again these differences were found only in the production of case errors: the 4-year-old and the 6-year-old groups did not differ in the amount of RCs lacking the correct case marking; however, the 4-year-old group produced more case errors than the adult group in both SRs ($z=-2.463$, $p<.05$) and in ORs ($z=-2.777$, $p<.05$).

The Wilcoxon test carried out for Experiment 3 revealed that in each of the groups the production of errors in SRs and ORs differed in the 5-year-old group ($z= -2.161$, $p<.05$) but not in the adult group ($z=-1.4$, $p=0.155$). Case errors were the only kind of error whose rate varied in SR and OR, but this happened only for children ($z= -3.79$, $p<.001$) and not for adults ($z= -0.68$, $p=.47$). In order to determine whether the types of errors differed between groups we carried out the Kruskal-Wallis test. The results revealed that the groups were different. Once again, these differences were observed only in case errors in ORs ($z=-3.73$, $p<.001$) but not in SRs ($z=-1.3$, $p=0.172$). The rest of errors were very infrequent (close to 0), and absent in the adult group.

5. Discussion

The main objective of the study was to investigate the production of RCs in Basque, a highly inflected, ergative SOV language with prenominal relatives. In previous comprehension experiments it had been established that unlike in nominative languages such as English, where SRs are generally comprehended with greater accuracy by both children and adults, in Basque, children and adults showed better performance with OR comprehension (Carreiras et al. 2010, Gutierrez-Mangado 2011). On the other hand, the results from the two production experiments reported here indicate that in production, better scores were obtained in SRs than in ORs, supporting the results obtained in nominative languages and contradicting the comprehension results mentioned for Basque. In what follows it will be shown that inverse pattern in the comprehension and production of RCs in Basque is fact only apparent. Moreover, we will defend that an explanation based on the visibility of case marking may offer a better account for the asymmetries reported in both child and adult performance.

In order to shed light on the SR vs. OR asymmetry, only RCs which could be unambiguously interpreted as SRs or ORs were included in the analysis. Ambiguity was observed in three kinds of structures: (i) RCs lacking an overt argument inside the RC, (ii) RCs with “ari” morphology and (iii) RCs with atypical word order. The lack of within-group differences in the ambiguous sentences eliminated in SRs and ORs led us to reject these productions as potential strategies used by the participants to avoid producing a specific type of RC. However, the between-group comparisons revealed a developmental pattern where 4 year-olds differed from adults in a) the production of more ORs lacking the overt S argument, b) the scarcity of “ari” type RCs and c) the more frequent production of RCs with atypical word order type. 5-year-olds differ from adults in the latter aspect whereas 6-year-olds show no differences with adults in the frequency of different kinds of ambiguous sentences in both SRs and ORs.

From the unambiguous RCs included for analysis, the percentage of grammatical RCs shows the opposite asymmetry observed in comprehension, namely, it was the case that for all groups in both production experiments the percentage of grammatical SRs was higher than for ORs. This difference was significant for the 4- and 5-year-old groups and for the adult group in Exp 2. The between-group comparison revealed, again, a developmental pattern as 4-year-olds show more errors than adults in both SRs and ORs, whilst 5-year olds only produce more errors in ORs and 6-year-olds do not differ significantly from adults.

If we look at the main source of error (45% before age 6 and below 20% afterwards), case errors seem to be responsible for the observed asymmetry, as observed in Graph 3. Importantly, few instances of resumptive pronouns, theta role or predicate change were found in both production experiments, which are reported among the typical avoidance strategies used in OR contexts in nominative languages (Labelle 1988, Belletti 2012) though their frequency seems to vary across languages as they are not frequent either in Spanish ORs (Ezeizabarrean, in press). The use of resumptive pronouns in our data did not differ in the children and adult groups and they were equally low in SRs and in ORs (<3% in all groups). As for the use of the passive, a strategy also reported to be used by children in order to avoid the production of an OR, it was only in the adult group of Exp 2 where it was found, suggesting perhaps that some adults were using more ‘formal’ language than the children during the task. The only error which was constant across all groups in both experiments was the case error.

Next we will try to argue that it is the visibility of the marking of the transitive S as ergative $-k$, which can account for this production/comprehension asymmetry. To start with, a clear piece of data emerged from Experiments 2 and 3, namely that the major source of error in the unambiguous RCs produced was case marking. Could it be the case that the ambiguous RCs produced and eliminated from the analysis were also case related? The answer seems to be affirmative. Let us take the omission of the internal argument of the RCs: this type of RC was produced specially by the 4-year-old group. One consequence of silencing the internal argument of the RC is that no case marking

needs to be done, that is, the RCs without an internal argument are in a way simpler than those with an overt argument, which need to be case marked. Though ergative case is attested very early in longitudinal spontaneous corpora, children do not use it in an adult-like way until age 3 and that it is one of the features which fossilize in child L2 acquisition (Ezeizabarrena, 2012). Recall that it is the subject argument inside the ORs which needs to be overtly case marked with the ergative, while the object argument in SRs bears absolutive (\emptyset) case marking. Moreover, as we saw in section 2, the production of “ari” RCs also affects directly the production of ergative case, as such structures involve case absorption (see (5) and (6)). The exact connection between these ambiguous structures and case marking needs to be further examined.

We suggest that the production/comprehension asymmetry reported could be explained by the ergative nature of Basque: ORs are favoured in comprehension and in processing, since the first argument in the RCs, which is overtly marked by the ergative case, is interpreted as the subject (agent) unambiguously. In contrast, in SRs the (zero marked) absolutive argument inside the RC seems to be interpreted by the children as either S (agent) or an O (patient). Notice, that the comprehension of SRs in Graph 1 reached values close to random. In production, most children’s target-deviant ORs are instances of RCs where the overt ergative case marking of the S argument has been omitted. This could be the result of some default strategy, according to which children overextend the interpretation of (zero-marked) absolutive arguments to S of transitive predicates, in addition to the typical interpretation of Os of transitive and Ss of intransitive predicates, in this ergative language. Such an interpretation is also consistent with the possibility of some allomorphic variation ($-k$, $-\emptyset$) in the case marking of Ss of transitive predicates attested in incomplete grammars (Ezeizabarrena 2012).

6. Conclusion

The production experiments reported here have shown that in general, Basque SRs are produced with greater accuracy than ORs. This result seems to contradict previous comprehension studies in Basque where ORs were reported to be more accurately comprehended. However, we have suggested that the overt case marking of the ergative case in Basque, together with the fact that the (ambiguous) absolutive case marking is zero-marked (and the same for intransitive S and O), may favour children’s and adults target-like comprehension of ORs. Thus, pronominal RCs starting with an overtly marked argument are identified as ORs from very early on, whereas RCs starting with a zero marked argument can be identified as either SRs or ORs. On the other hand, in production, SRs are produced with higher accuracy and interestingly the most common error in ORs is the absence of the ergative case marking. Therefore, we conclude that the ergative-absolutive case system plays a crucial role in the asymmetries observed in production and comprehension of RCs in Basque.

References

- Adani, Flavia, Heather K.J. van der Lely, Matteo Forgiarini & Maria Teresa Guasti. 2010. Grammatical feature dissimilarities make relative clauses easier: A comprehension study with Italian children. *Lingua* 120. 2148-2166.
- Altube, Severo. 1929. *Erderismos*. Bermeo. 2nd edn. Bilbao 1975.
- Arosio, F. Adani, Flavia and Guasti, Maria Teresa. 2005. Processing gramatical features by Italian children. In A. Belletti, E. Bennati, C. Chesi and I. Ferrari (eds.), *Acquisition and development. Proceedings of GALA 2005* (pp15-27). Siena: Cambridge Scholars Press.
- Artiagoitia, Xabier. 1992. Why Basque doesn't relativize everything. *International Journal of Basque Linguistics and Philology* XXVII, Donostia: Gipuzkoako Foru Aldundia. 11-35.
- Belletti, Adriana. 2012. Considering the complexity of relative clauses and passives from the Italian perspective. In Sandrine Ferré, Philippe Prévost, Laurie Tuller and Rasha Zebib (eds.) *Proceedings of the Romance Turn IV workshop on the acquisition of Romance Languages*. Cambridge Scholars Publishing.
- Bever, Thomas G. 1970. The cognitive basis for linguistic structure. In John. R. Hayes (ed.), *Cognition and the development of language*, 279-362. New York: Wiley.
- Brown, H. Douglas. 1972. Children's comprehension of relativized English sentences. *Child Development* 42. 1923-1936.
- Carreiras, Manuel, Jon Andoni Duñabeitia, Marta Vergara, Irene de la Cruz & Itziar Laka. (2010). Subject relative clauses are not universally easier to process: Evidence from Basque. *Cognition* 115. 79-92.
- Chang, Hsing-Wu. 1984. The comprehension of complex Chinese sentences by children: Relative clause. *Chinese Journal of Psychology* 26. 57-66.
- Cheng, Sherry Yan-Yin. 1995. *The acquisition of relative clauses in Chinese*. Master's Thesis, National Taiwan Normal University.
- Cho, Sookeun. 1999. *The acquisition of relative clauses: Experimental studies in Korean*. University of Hawai'i at Manoa unpublished dissertation.
- Correa, Leticia M. 1995. An alternative assessment of children's comprehension of relative clauses. *Journal of Psycholinguistic Research* 24. 183-203.
- Courtney, Ellen H. 2006. Adult and child production of Quechua relative clauses. *First Language* 26(3). 317-338.
- de Rijk, Rudolf. 1969. Is Basque an SOV language? *FLVI*.
- de Villiers, Jill, Peter de Villiers & Esme Hoban. 1994. The central problem of functional categories in English syntax of oral deaf children. In Helen Tager-Flusberg (ed.), *Constraints on Language Acquisition: Studies of Atypical Children*, 9-47. Hillsdale, NJ: Erlbaum.
- Ezeizabarrena, M. J. 2012. The acquisition of the (in)consistent ergative marking in Basque: L1 and early L2. In I. Laka & B. Fernández (eds). Special Issue accounting for ergative languages. *Lingua* 122, (3), 303-317.
- Ezeizabarrena, M.J. (in press). Children do not substitute object relatives with subject relatives in every romance language: the case of Spanish. In Avram (ed.) *Revue Romaine de Linguistique*.

- Friedmann, Naama. 2010. The production of relative clauses by 5 year olds across multiple languages: they prefer to be the children who do not produce object relatives. Paper presented at the Cost A-33 Conference: *Let the children speak. Learning of critical language skills across 28 languages*. London.
- Friedmann, Naama & Rama Novogrodsky. 2004. The acquisition of relative clause comprehension in Hebrew: A study of SLI and normal development, *Journal of Child Language* 31. 661-681.
- Friedmann, Naama, Adriana Belletti & Luigi Rizzi. 2009. Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua* 119. 67-88.
- Grodzinsky, Yosef. 1989. Agrammatic comprehension of relative clauses. *Brain and Language* 37. 480-499.
- Gutierrez-Mangado, M. Juncal. 2011. Children's comprehension of relative clauses in an ergative language: the case of Basque. *Language Acquisition*, 18 (3), 176-201.
- Gutierrez-Mangado, M. Juncal & Mari Jose Ezeizabarrena. 2010. *Basque children produce subject relatives with more accuracy than they comprehend them*. Poster presented at the Romance Turn IV, The Université François-Rabelais, Tours, France. 25-27 August, 2010.
- Håkansson, Gisela & Kristina Hansson. 2000. Comprehension and production of relative clauses: A comparison between Swedish impaired and unimpaired children. *Journal of Child Language* 27. 313-333.
- Hakuta, Kenji. 1981. Grammatical description versus configurational arrangement in language acquisition. *Cognition* 9. 197-236.
- Hamburger, Henry & Stephen Crain. 1982. Relative acquisition. In Stan Kuczaj (ed.), *Language Development, vol. 1, Syntax and Semantics*. Hillsdale, NJ: Lawrence Erlbaum.
- Hsu, Chun-Chieh Natalie, Gabriella Hermon & Andrea Zukowski. 2009. Young children's production of head-final relative clauses: Elicited production data from Chinese children. *Journal of East Asian Linguistics* 18. 323-360.
- Hualde, Jose Ignacio & Jon Ortiz de Urbina. 1987. Restructuring with ARI. *ASJU International Journal of Basque Philology*, XXI-2, 425-452.
- Kanno, Kazue. 2007. Factors affecting the processing of Japanese relative clauses by L2 learners. *Studies in Second Language Acquisition* 29. 197-218.
- Labelle, Marie. 1988. *Prédication et mouvement: le développement de la construction relative chez les enfants francophones*, unpublished doctoral dissertation, University of Ottawa, Canada.
- Laka, Itziar. 2006. On the nature of Case in Basque: structural or inherent? In Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz and Jan Koster (eds.) *Organizing Grammar*, 374-382. Berlin: Mouton de Gruyter.
- Lee, Kwee-Ock. 1991. *On the first language acquisition of relative clauses in Korean: The universal structure COMP*. Ithaca, NY: Cornell University unpublished dissertation.
- Lee, Thomas Hun-Tak. 1992. The inadequacy of processing heuristics: Evidence from relative clause acquisition of Mandarin Chinese. In Thomas Hun-Tak Lee (ed.) *Research on Chinese Linguistics in Hong Kong*, 47-85. Hong Kong: Linguistics Society of Hong Kong.
- Mak, Willem M., Wietske Vonk & Herbert Schriefers. 2002. The influence of animacy on relative clause processing. *Journal of Memory and Language* 47. 50-68.

- McKee, Cecile, Dana McDaniel & Jesse Snedeker. 1998. Relatives children say. *Journal of Psycholinguistic Research* 27. 573-596.
- Ortiz de Urbina, Jon. 1990. *Parameters in the Grammar of Basque*. Dordrecht: Foris.
- Oyharçabal, Beñat. 1987. *Etude descriptive de constructions complexes en basque: propositions relative, temporelles, conditionnelles et concessives*. University of Paris 7 dissertation.
- Ozeki, Hiromi & Yashuhiro Shirai. 2005. Semantic bias in the acquisition of relative clauses in Japanese. Paper presented at the 29th annual Boston University Conference on Language Development. Boston, MA.
- Ozeki, Hiromi & Yashuhiro Shirai. 2007. The consequence of variation in the acquisition of relative clauses: An analysis of longitudinal production data from five Japanese children. In Yoshiko Matsumoto, David Y. Oshima, Orrin R. Robinson & Peter Sells (eds.). *Diversity in languages: Perspectives and implications*. 243-270. Stanford, CA: CSLI Publications.
- Salaburu, Pello. 1987. *Euskal morfositaxia eta fonologia: eztabaida gaiak*. Bilbo: Euskal Herriko Unibertsitatea.
- Sauerland, Uli & Edward Gibson. 1998. How to predict the relative clause attachment preference. Paper presented at the 11th CUNY sentence processing conference, Rutgers University, New Brunswick, NJ.
- Stavroulaki, Stavroula. 2001. Comprehension of reversible relative clauses in specifically language impaired and normally developing Greek children. *Brain and Language* 77. 419-431.
- Su, Yi Ching. 2004. Relatives in Mandarin children. Paper presented at Generative Approaches to Language Acquisition in North America, University of Hawai'i at Manoa.
- Traxler, Matthew J., Robin K. Morris & Rachel E. Seely. 2002. Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language* 47. 69-70.
- van der Lely, Heather. 1996. Specifically language impaired and normally developing children: Verbal passive vs adjectival passive sentence interpretation. *Lingua* 98. 243-272.
- Yip, Virginia & Stephen Matthews. 2007. Relative clauses in Cantonese-English bilingual children. *Studies in Second Language Acquisition* 29. 277-300.