

Finding identity in the midst of ambiguity: Case and number disambiguation in Basque

Bojana Ristic^{1*}, Simona Mancini¹ & Nicola Molinaro^{1,2}

¹ *BCBL – Basque Center on Cognition, Brain and Language, San Sebastián, Spain;* ² *Ikerbasque, Bilbao, Spain*

*Bojana Ristic, Paseo Mikeletegi 69 (BCBL), 20009 San Sebastián, Spain. Phone: +34 943 309 300 (Ext. 314). Email: b.ristic@bcbl.eu

Finding identity in the midst of ambiguity: Case and number disambiguation in Basque

Restrictive contextual information has been found to bias syntactic disambiguation, when only one alternative leads to a meaningful interpretation. The current study tests whether disambiguation can be influenced by *nonrestrictive* cues – when several alternatives are equally plausible. We first evaluated if modifier number biased the disambiguation of number- and case-ambiguous nouns in Basque. In a noun phrase comprehension paradigm, ambiguous noun number judgments were biased by preceding modifier number. Then, using a preamble completion paradigm, we examined whether headnoun disambiguation and thus sentence completion was also biased by modifier number. Our results suggest that *nonrestrictive* information (singular and plural number) can affect disambiguation. We also report task differences in the overall interpretation of ambiguous Basque nouns, as well modifier-induced agreement errors. We suggest that the parser uses *any* available context information when there is ambiguity, including preceding modifier markings.

Keywords: subject-verb agreement, morphological ambiguity, Basque, number and case attraction

1.1. Syntactic ambiguity and disambiguating information

Syntactic ambiguity has received a lot of attention in psycholinguistic research. It has proved to be a useful tool for investigating sentence comprehension and human sentence processing because it reveals parsing preferences. Sentence processing models that have emerged from this line of research can broadly be divided into two streams, distinguished primarily on the basis of the time-course of ambiguity resolution. *Serial* models assume that only one syntactic alternative is initially available (e.g. Frazier & Rayner, 1982). *Parallel* or *constraint-based* models, however, assume that several alternatives are available in parallel at all stages of processing, and that these alternatives are continuously evaluated based on different constraints (e.g. MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, 1996).

Despite these differences related to timing and constraints, both types of models acknowledge the impact of *relevant linguistic and non-linguistic constraints*, or *syntactically relevant contextual information*, on syntactic ambiguity resolution (Spivey, Tanenhaus, Eberhard, & Sedivy, 2002; Trueswell, 1996). These constraints, among others, might include: animacy of the subject noun phrase (Trueswell, Tanenhaus, & Garney, 1994); semantics of preceding words (Ni & Crain, 1990; Tabossi, Spivey-Knowlton, McRae, & Tanenhaus, 1994) or the preceding sentence (Marslen-Wilson & Tyler, 1987); referential context (Altman & Steedman, 1988); frequency and lexical features of preceding verbs (MacDonald, 1994; Malaia, Wilbur, & Weber-Fox, 2010; Trueswell, Tanenhaus, & Kello, 1993), as well as their tense (Trueswell & Tanenhaus, 1991). Several studies have also found that prosody triggers disambiguation in the case of words that are ambiguous between different arguments. In these studies, mostly conducted on German, syntactic ambiguity is generated by using syncretic case markers (Knoeferle, Crocker, Scheepers, & Pickering, 2005; Kröger, Münster, & Knoeferle, 2017; Weber, Grice, & Crocker, 2006). For example, in a visual world paradigm, Weber et al. (2006) studied temporarily structurally ambiguous sentences in German, in which the initial noun was ambiguous between a subject in nominative case and an object in accusative case (e.g. *Die Katze...*, “the cat[NOM/ACC]...”), and disambiguated at the second noun in the sentence (*Die Katze[NOM] jagt womöglich den Vogel[ACC]*, “The cat is possibly chasing the bird”, or *Die Katze[ACC] jagt*

womöglich der Hund[NOM], “The cat is possibly chased by the dog”). The authors found that prosody could bias the interpretation of the ambiguous word and influence the general preference towards a nominative case reading even before the second noun has been encountered.

Importantly, in most studies to date, disambiguating contextual information restricts the parser to one of the possible interpretations. Indeed, not selecting this interpretation would entail choosing an interpretation that is implausible, semantically (or prosodically) incongruent, or even ungrammatical. We will call this type of disambiguating contextual information *restricting*. By contrast, *nonrestricting* contextual information refers to cases in which all possible interpretations are equally viable. When *nonrestricting* cues are present, the parser is not restricted to one alternative, i.e. there is no choice that should be preferred, or must be selected in order to produce a congruent grammatical sentence.

It is still an open question whether the parser uses *nonrestrictive* contextual information in syntactic disambiguation. The aim of the current study was to answer this question. More specifically, we investigated whether syntactic disambiguation was influenced merely by number marking on a noun phrase (NP) modifier, even though this would not determine parsing options in any way, and therefore would not guide language users to a specific interpretation of the sentence. To this end, we used ambiguously marked nouns, preceded by number-marked modifiers. In Basque, nouns marked with *-ak* ambiguously indicate either absolutive case and plural number, or ergative case and singular number. Unlike materials used in previous studies, this Basque syncretism creates ambiguity that entails not only subject/object ambiguity, but also transitive/intransitive construction ambiguity. Thus, the present study tests whether number marking on an element in the immediate context (a modifier) can bias the resolution of temporary structural ambiguity, even though not complying with this bias would have no consequences for processing.

To our knowledge, the only studies that have looked at whether the parser uses a neighboring element’s morphological information (i.e. *nonrestrictive* information) to resolve syntactic ambiguity have addressed case ambiguity in German (Bader, 1997; Bader & Meng, 1999). However, these

studies focused on how resolving temporary ambiguity using neighboring information caused erroneous interpretations and hindered sentence comprehension downstream. In the current study, using or not using the available contextual information has no consequences for the acceptability of the resulting sentences. Furthermore, most of the abovementioned studies investigated syntactic ambiguity in sentence comprehension, either measuring reading times related to the preferred vs. dispreferred analysis, or examining behavior when participants listened to temporarily ambiguous sentences. In the current study, we implemented a number judgement paradigm and a sentence production paradigm. Firstly, in Experiment 1, we tested whether the ambiguous NPs in isolation were interpreted as singular ergative or plural absolutive verb forms, and, more importantly, whether the contextual information (the number of the modifier) affected that decision. Then, in Experiment 2, we presented the same NPs as preambles in a sentence completion task, to test whether contextual information was used for syntactic disambiguation in sentence production. The following section describes the features of Basque NPs that are relevant for the current study.

1.2. Basque NPs and morphological ambiguity

In Basque NP structure, modifiers marked with a genitive morpheme must precede head nouns, as illustrated in (1).

- (1) mutilaren laguna
boy-det-GEN friend-det
“the boy’s friend”

The genitive marker *-ren* can be used to express either possession, agency, or theme (Laka, 1996). In this experiment, only the relation of possession is investigated, so we will refer to the nominal modifier marked with the *-ren* morpheme as a “possessive (genitive) modifier”. Furthermore, subjects of transitive verbs in Basque are marked by ergative case (*-(a)k* for singular, *-ek*

for plural), while subjects of intransitive verbs are marked by absolutive case (zero morpheme, thus undistinguishable from the determiner, *-a* for singular, *-ak* for plural). This means that the nominal suffixes *-ek* (ergative plural) and *-a* (absolutive singular) are unambiguous, while the *-ak* sequence is ambiguous. The latter is therefore employed in the current study. The *-ak* sequence denotes either singular number and ergative case (a combination of the singular determiner morpheme *-a* and the ergative marker *-k*), as shown in (2a), or plural number and absolutive case (the plural determiner morpheme *-ak* and zero morpheme), as shown in (2b). Additionally, in transitive constructions, objects take the absolutive case (plural or singular), as shown in (2c). The sequence *-ak* can therefore be used to mark singular subjects of transitive verbs, plural subjects of intransitive verbs or plural objects of transitive verbs.

(2) a) *Mutilak_{ERG} ogia erosi du*

boy-det-ERG.SG bread-det-ABS.SG buy has

“The boy has bought a bread”

b) *Mutilak_{ABS} etorri dira*

boy-det-ABS.PL arrive are

“The boys have arrived”

c) *Gizonak mutilak_{ABS} ikusi ditu*

man-det-ERG.SG boy-det-ABS.PL see has

“The man has seen the boys”

Importantly, previous research has shown that an object reading of sentence-initial nouns is highly disfavoured due to SOV word order preference (Erdocia, Laka, Mestres-Missé, & Rodriguez-Fornells, 2009; Erdocia, Zawiszewski, & Laka, 2014)¹. Therefore, we will disregard the object reading in the rest of the paper and focus only on the ergative singular/absolutive plural subject ambiguity.

Using a visual world paradigm, Yetano, Duñabeitia, & Laka (2011) looked at the processing of spoken sentences containing an initial animate NP argument marked by *-ak*. They found higher numbers of refixations on objects referenced by the initial NP, when the verbs were disambiguated as having a plural object/intransitive subject (absolutive case), compared to when they were disambiguated as having a singular subject (ergative case). Since refixations are usually interpreted as part of a reanalysis process, the authors conclude that Basque speakers prefer to interpret an ambiguous head noun as taking a singular subject (ergative). They suggest that speakers interpret initial ambiguous animate head nouns as the higher-ranking argument of the event due to an agent-initial processing preference.

In the current study, we investigated whether ambiguously marked nouns in Basque could be disambiguated by *nonrestrictive* information, more specifically, by the number of their modifier. If this was the case, we expected that nouns marked by *-ak* would be interpreted as ergative singular when preceded by a singular modifier, and as absolutive plural when preceded by a plural modifier. Moreover, we tested this question within two different configurations: first, when the ambiguous NPs were presented in isolation, and then, in a sentence production paradigm. Our experiments complement existing findings on ambiguous NPs in Basque, which provided results using sentence comprehension paradigms.

2. Experiment 1

Experiment 1 focuses on the processing of ambiguously marked nouns and their modifiers in isolation, i.e. at the phrase comprehension level. We presented participants with NPs containing a number-ambiguous head noun and either a singular or plural possessive genitive modifier, and asked them to judge whether the ambiguous noun in the NP was plural or singular. In Basque, this type of modifier precedes the head noun. We manipulated its number (singular/plural), in order to see if this would influence the reading of the head noun. Therefore, more ergative singular

interpretations of the ambiguous NP are expected in the singular condition, and more absolutive plural interpretations in the plural condition.

A previous study by Yetano et al. (2011) showed that the ergative singular interpretation is overall preferred in sentence processing. However, it was unclear whether this would still be the case if NPs were presented in isolation, as in the current experiment. Importantly, if an element modifying a number-ambiguous head noun does indeed have some influence on disambiguation, its number marking should modulate this potential preference. That is, if the ergative singular interpretation of the ambiguous NP is overall preferred in isolation, this preference should be stronger in the singular condition; conversely, if the absolutive plural interpretation is preferred, the plural preference should be stronger in the plural condition.

2.1 Methods

2.1.1. Participants

Twenty four Basque-Spanish bilinguals took part in the experiment (11 male; mean age = 29.03, SD = 5.72). Their dominant language was Basque, which they had acquired within the first 3 years of life. Their scores on the Basque proficiency test were high (BEST² mean = 63.75 out of 65, SD = 1.50). The experiment was approved by the BCBL Ethics Review Board and complied with the guidelines of the Helsinki Declaration.

2.1.2. Materials

The materials for Experiment 1 consisted of 40 NPs containing ambiguous head nouns (see Table 1). Each NP consisted of two nouns – a possessive genitive modifier in either the singular or plural form, followed by the head noun it modified. The modifier was always marked by the possessive marker *-(r)en*. The ambiguous head nouns were always marked by *-ak*. Importantly, the modifier was *nonrestrictive*, i.e. interpreting *nebak* (Table 1) as either “brother” or “brothers” is

equally possible with both singular modifier (*gaixoaren*, “of the patient”) or plural modifier (*gaixoen*, “of the patients”). Also, NPs followed specific constraints: both nouns were animate; nouns denoting occupations were avoided; nouns whose singular/plural reading might be preferred were avoided (*parents*, *wife*, etc.).

[Insert Table 1 here]

To test whether certain combinations of singular/plural modifiers and singular/plural head nouns resulted in more natural NPs, we performed a naturalness judgement task with 20 participants who did not take part in either Experiment 1 or 2. In this task, the experimental items occurred in 4 forms: in their unambiguously plural form (ergative plural) with singular and plural genitive modifiers (*emakumearen* [GEN.SG]/ *emakumeen* [GEN.PL] *adiskideek* [PL.ERG], ‘the woman’s/women’s friends’), and in their unambiguously singular form (absolutive singular) with singular and plural genitive modifiers (*emakumearen* [GEN.SG]/ *emakumeen* [GEN.PL] *adiskidea* [SG.ABS], ‘the woman’s/women’s friend’). The items were Latin Squared and rated on a 1-5 point scale (1 = completely unnatural; 5 = completely natural). High naturalness scores were obtained in all conditions, with an average score of 3.8 ($SE = 0.05$; means per condition: ergative plural/singular modifier: 3.92, $SE = 0.10$; ergative plural/plural modifier: 3.57, $SE = 0.10$; absolutive singular/singular modifier: 4.01, $SE = 0.09$; absolutive singular/plural modifier: 3.68, $SE = 0.10$). Furthermore, we performed linear mixed effects models analysis to examine potential differences in naturalness across the conditions. The factors were dummy-coded (singular = 0, plural = 1). The best fitting and maximally converging model included the Modifier number X Head noun number interaction as a fixed effect, random by-item and by-subject random intercepts, as well as random by-item and by-subject random slopes for Modifier number and Head noun number. We found a significant effect of Modifier number, such that NPs with singular modifiers were rated as more natural (Intercept (Plural condition): -0.17, Estimate: 0.26, SE: 0.11, $t = 2.34$). No other effects were significant (all $ts < 1$). Furthermore, the frequency of the genitive modifier nouns as they appeared in the experiment, e.g. in their singular and plural forms, was checked using the Basque Frequency

Dictionary (Acha, Laka, Landa, & Salaburu, 2014). One noun couldn't be found in the dictionary. The paired t-test on the remaining 39 genitive modifiers showed no significant difference between singular and plural forms ($t = -1.68$, $df = 38$, $p = 0.10$). Importantly, the lack of the significant interaction between the Modifier number and Head noun number suggests that the certain combinations of singular/plural modifiers and singular/plural head nouns do not result in more natural NPs, testifying that the modifiers are indeed *nonrestrictive*.

The experimental items were divided into two lists. Eighty unambiguous filler items were then added to each list. Half of the fillers contained unambiguously singular (absolute singular, the suffix *-a*) and the other half contained unambiguously plural (ergative plural, the suffix *-ek*) head nouns. The overall amount of singular and plural information on both head nouns and modifiers was thus balanced.

The NPs were followed by a question regarding the number-ambiguous head noun (e.g. *How many boys are there?*³), with two possible answers (*one* or *more than one*). Moreover, half of the NPs were followed by a question regarding the unambiguous head noun, and the other half by a question regarding the modifier. This was done in order to make sure that the participants would pay attention to both the modifier and the head noun.

2.1.3. Procedure

The NPs were presented using PsychoPy software, version 1.83.04 (Peirce et al., 2019). They were presented in the centre of the screen for 2 seconds, after which a screen appeared with a question in the centre and the two answer options appeared at the bottom left and bottom right. The positions of each option and the positions of the correct answers were balanced across the experiment. The maximum time allowed for an answer was 5 seconds. The experiment consisted of five practice trials and 120 experimental trials, presented in two blocks. The participants were tested individually in a sound-proof cabin and the whole session lasted approximately 15 minutes.

2.1.4. Data Analysis

The participants' answers were coded for singular or plural answers (singular = 0, plural = 1). Generalised linear mixed effects models with binomial family were fitted with the singular/plural answer as a dependent variable, and the Type of Modifier (Singular, Plural) as the fixed effect. We tested a series of models for exclusion of random by-item and by-subject slopes (by-item and by-subject random intercepts were included by default), using a forward-direction model selection method and the "best-path" algorithm (Barr, Levy, Scheepers, & Tily, 2013). Slopes were included if they fell below the model-selection α level of .05. When encountering non-convergence problems, the random effects structure was progressively simplified (by removing random slopes – by-item and by-participant independently) until convergence was reached (Barr et al., 2013). We report the intercept, the estimate, standard error, z and p values. All the analyses were done using the package *lme4* (Bates, Machler, Bolker, & Walker, 2015) in R Statistical software (Team, 2014). In order to evaluate the effect of the singular and plural modifiers in more detail, we also performed two post-hoc Chi-Square tests on the Singular and Plural conditions separately. Another Chi-Square test was performed on the full dataset to test for a possible overall preference for using the ergative singular/absolute plural. The Bonferroni correction for the multiple comparisons was performed using the *p.adjust* function from the *stats* package in R (R Core Team, 2017).

Although the filler items did not constitute experimental conditions and thus were not Latin Squared, the performance on these items was additionally analyzed in order to get more insight into whether the effects were found with unambiguous phrases, or indeed triggered by the need to resolve ambiguity. For this analysis, we fitted a model with accuracy as a dependent measure, and the Type of Modifier (Singular, Plural) x Headnoun number (Singular, Plural) interaction as fixed effects. Model selection and the rest of the procedure followed that used for the experimental items.

2.2. Results and discussion

The best fitting (and maximally converging) model included the fixed effect Type of Modifier, as well as by-item and by-subject random intercepts and slopes. We found a significant effect of Type of Modifier, such that plural answers were used significantly less in the Singular than in the Plural modifier condition [Intercept (Plural condition): 1.62, Estimate: -0.52, SE: 0.22, $z = -2.28$, $p = .02$].⁴ Figure 1 shows the use of plurals per condition, Figure 2 shows the number of singular/plural answers per condition, and Table 2 shows the proportions of singular and plural answers per condition. The Chi-Square tests showed a significant difference between singular and plural answers in the Singular ($X\text{-squared} = 83.37$, $df = 1$, $p < .001$, adjusted $p < .001$), as well as in the Plural condition ($X\text{-squared} = 127.44$, $df = 1$, $p < .001$, adjusted $p < .001$), in both cases driven by the increase in plural answers. As can also be seen in Figure 2, plural answers were comparably dominant in both Singular and Plural conditions, but their proportion increased by 6.23% in the Plural condition. At the same time, the proportion of singular answers in the Singular condition increased by 22.12%. Thus (although not significant, probably due to the dominance of plural answers), the singular modifier also triggered an increase in singular answers in the Singular condition, leading to the effect of Type of Modifier.

We also found an overall increase in plural answers ($X\text{-squared} = 208.39$, $df = 1$, $p < .001$, adjusted $p < .001$). This overall preference for the plural answers goes against an earlier finding in the literature on sentence comprehension (Yetano et al., 2011), where a preference towards an ergative singular reading was observed. We thus present novel evidence that NPs in Basque ambiguously marked by *-ak* are preferably interpreted as (absolute) plural when presented in isolation. This is further elaborated in the General discussion. Since the absolute plural preference was found to be modulated by modifier number (the effect of Type of Modifier), our results imply that the number marking of a preceding word is used as a *nonrestrictive* cue to disambiguate syncretic forms and thus resolve the ambiguity of the NP.

Participants showed an accuracy rate of 86% ($SE = 0.008$) on filler items. The best fitting model included the Type of Modifier (Singular, Plural) x Headnoun number (Singular, Plural) interaction as a fixed effect, as well as by-item and by-subject random intercepts. We found a significant effect of Type of Modifier, such that the singular modifier condition elicited higher

accuracy overall [Intercept (Plural condition): 1.73, Estimate: 0.71, SE: 0.28, $z = 2.54$, $p = .01$], suggesting that plural modifiers (either with singular or plural headnouns) trigger more errors. However, no main effect of the Headnoun number, nor Type of Modifier x Headnoun number interaction was found. Although the current study was not designed to compare ambiguous and unambiguous headnouns directly, this additional analysis suggests that headnouns unambiguously marked as singular (-*a*) or plural (-*ek*) might elicit comparable performance, and that performance is not affected by the number of the modifier, as was the case with the experimental items. It therefore seems ambiguity is a necessary condition for the modifier to affect answer choice. It appears that *nonrestrictive* number cues are only used when it becomes necessary, i.e. in the context of morphosyntactic ambiguity.

[Insert Table 2 here]

[Insert Figure 1 here]

[Insert Figure 2 here]

3. Experiment 2

After testing whether *nonrestrictive* cues biased the disambiguation of ambiguous NPs, we proceeded to test the same NPs in a sentence production context. In Experiment 2, we tested whether *nonrestrictive* information (i.e. the number of the preceding modifying noun) could bias interpretation of an ambiguous sentence-initial head noun, and consequently the completion of the sentence.

We presented participants with preambles (NPs from Experiment 1) and asked them to freely complete the sentences as they liked. We wanted to see if the presence of singular/plural-marked modifier could bias the reading of the ambiguous head noun towards a singular or plural interpretation, and consequently the choice of the verb towards intransitive (and thus plural absolutive verb agreement), or transitive (and thus singular ergative verb agreement; see Section 1.2).

Based on the previous study by Yetano et al. (2011), we expected the ergative singular interpretation to be overall preferred in sentence processing. Moreover, if the element modifying the number-ambiguous head noun indeed influences morphosyntactic disambiguation, this ergative singular preference should be modulated. More specifically, we expected genitive singular modifiers to boost the ergative/singular reading of *-ak* head nouns and, consequently, the choice of transitive verbs. In contrast, plural genitive modifiers should reduce ergative singular preference by increasing the plural/absolutive interpretation of ambiguous NPs and the subsequent choice of intransitive verbs.

3.1. Methods

3.1.1. Materials

The materials for Experiment 2 consisted of 40 preambles, which were the same NPs as those used in Experiment 1. The items were recorded by a native Basque speaker in standard Basque.

3.1.2. Participants

30 native speakers of Basque (10 male; mean age = 22.8, SD = 3.08) took part in Experiment 2, none of whom had participated in Experiment 1, nor in the naturalness norming. All of them were bilinguals (17 Basque-dominant, 13 Spanish-dominant) with high scores on the Basque proficiency test (BEST mean = 62.67 out of 65, SD = 2.03), and had acquired Basque within the first 3 years of life. The experiment was approved by the BCBL Ethics Review Board and complied with the guidelines of the Helsinki Declaration.

3.1.3. Procedure

Participants were tested individually in a sound-proof cabin. The preambles were presented auditorily and participants were required to repeat the preamble and complete the sentence orally to

form a grammatical sentence in Basque. They were instructed to complete the sentence as they wished, and to answer intuitively. Sentence production was recorded using a headset microphone. The experiment consisted of five practice trials and 120 experimental trials, presented in two blocks.

3.1.4. Data analysis

Before analyzing the answers for the ambiguous NPs, 4 native Basque speakers transcribed them⁵, then coded the use of the ergative singular/absolute plural, as well as any errors. The choice of the best model and the rest of the data analysis were based on the same procedures as those in Experiment 1.

3.2. Results and discussion

78.75% of answers on experimental trials were correct, while 16.16% were errors; no answers were given on the remaining 5.08% trials.

As regards correct answers, 65% (SE = 0.02) in the Singular genitive condition, and 52% (SE = 0.02) in the Plural genitive condition were ergatives. The best fitting (and maximally converging) model included the fixed effect of condition and by-item and by-subject random intercepts and slopes. We found a significant effect of Type of Modifier, such that the ergative singular was used significantly more in the Singular than in the Plural modifier condition [Intercept (Plural condition): 0.08, Estimate: 0.89, SE: 0.32, $z = 2.75$, $p = .006$].⁶ Figure 3 shows the use of the ergative singular per condition, while Figure 4 shows the use of both cases per condition. The Chi-Square tests for Singular condition showed a significant increase in the use of the ergative singular (X-squared = 44.31, $df = 1$, $p < .001$, adjusted $p < .001$). In the Plural condition, no significant difference was found (X-squared = 0.58, $df = 1$, $p = 0.45$, adjusted $p = 1$), but the use of the plural absolute did increase by 21.02% (see Figure 4). This resulted in a comparable use of the absolute plural and (the overall

preferred) ergative singular in this condition. This finding implies that plural modifiers also influence number interpretation for ambiguous nouns.

The Chi-Square tests demonstrated an overall preference for ergative constructions (X-squared = 28.81, $df = 1$, $p < .001$, adjusted $p < .001$). This is in contrast to the results in Experiment 1, but in line with previous research on the processing of sentence-initial ambiguously marked nouns in Basque (Yetano et al., 2011), and will be elaborated in the General discussion. Altogether, the results indicate that the disambiguation of a number/case-ambiguous sentence-initial NP can be influenced by *nonrestrictive* contextual information (number of the head noun's modifier), in line with the results of Experiment 1.

[Insert Figure 3 here]

[Insert Figure 4 here]

Given that errors constituted a significant percentage of our data, we performed an additional analysis of those trials in which participants used verb and number agreement resulting in an ungrammatical sentence. For each erroneous trial, the error type was coded based on the nominal suffix, i.e. based on the nominal suffix that would have been correct given the continuation (i.e. auxiliary) produced. Therefore, the possible errors for experimental items were: *-ak* used for *-a* (using a noun marked by *-ak* with an auxiliary that required the absolutive singular, a singular error), or *-ak* used for *-ek* (using a noun marked by *-ak* with an auxiliary that required an ergative plural, a plural error). That is, the singular error meant that the ergative singular/absolutive plural marked noun was substituted for an absolutive singular, while the plural error meant it was substituted for an ergative plural (see Section 1.2). The error type was used as a dependent variable in the analysis, with the “*-ak* for *-a*” or singular error coded as 0, and “*-ak* for *-ek*” or plural error coded as 1. We then fitted generalised linear mixed-effect models with binomial family, with Type of Modifier (Singular, Plural) as the fixed effect, following the procedure described for the correct trials, also using the Chi-Square test in the same way. The maximal best fitting model included Type of Modifier as a fixed effect and a by-participant random intercept. We found a significant effect of the Type of Modifier [Intercept

(Plural condition): 5.40, Estimate: -2.82, SE: 0.92, $z = -3.05$, $p = .002$], such that the “-ak for -ek” or plural error was produced significantly less in the Singular modifier condition. Figure 5 shows the distribution of errors per condition. We found that 90.72% of all errors in experimental items were of the “-ak for -ek” type (i.e. plural errors), while singular errors (“-ak for -a”) constituted only 8.42% of the total number of errors. The remaining 1.05% of errors belonged to other types (e.g. omission, unintelligible completion, etc) and were not included in the analysis. It is worth noting that the dominant type of errors we found, i.e. mixing -ek and -ak markings, seems to be rather common in spoken Basque, as reported by native speakers. Not surprisingly, additional analyses showed that plural errors occurred significantly more in both the Plural condition (X-squared = 100.15, $df = 1$, $p < .001$, adjusted $p < .001$), and the Singular condition (X-squared = 33.8, $df = 1$, $p < .001$, adjusted $p < .001$). Furthermore, Figure 5 shows that singular errors increased in the singular condition, and plural errors increased in the plural condition. Additionally, a comparable number of errors was produced by Spanish and Basque dominant participants (Spanish dominant: 44.15 %, Basque dominant: 55.85 %), and including Language Dominance as a factor did not improve the model fit, nor did it generate a significant effect ($z = -0.12$) or interaction with the Type of Modifier factor ($z = 0.30$).

[Insert Figure 5 here]

In sum, the error analysis showed that plural errors were overall more frequent, and that the presence of a plural modifier led to more plural number agreement errors. Additionally, although singular errors were rare, we observed that they also increased in the singular modifier condition. The fact that the modifier’s feature spilled over to the entire NP incorrectly influencing subsequent agreement reflects the *number attraction* phenomenon. Here, the number of the subject head noun modifier (*attractor*) affects processing of head noun number, thus causing subject-verb number agreement errors (**The key to the cabinets were*, Bock & Miller, 1991; Eberhard, 1997, among many other examples). In the current experiment, we saw that, in conjunction with the overall dominance of plural errors, both plural and singular modifiers acted as *attractors*.

Although the filler items were not Latin Squared as they did not constitute experimental conditions, as in Experiment 1, we additionally analyzed them to obtain more information regarding the performance in the absence of ambiguity. Average accuracy for filler items was 85.66 % (SD = 0.35); 11% of answers were errors, and 3.34% were missing or incomplete answers. The dominant type of error was a transitivity error, where the absolutive singular noun was used as an ergative singular (“-a for -ak[ERG.SG]”, 36.16% of all errors) and where the ergative plural was used as an absolutive plural (“-ek for -ak[ABS.PL]”, 49.82% of all errors). Importantly, number errors constituted only 10.70% of all errors, and a Chi-Square test on these errors showed no difference between the conditions (X-squared = 0.31, df = 1, p = 0.58).

The results of Experiment 2 show that modifier marking caused comparable error rates (attraction errors) and disambiguation effects: singular modifiers triggered more singular ergative use and increased singular errors, while plural modifiers triggered more plural absolutive use and increased plural errors. However, the number of the modifier did not affect the answers in the unambiguous filler items, supporting our finding from Experiment 1 that number marking is used as a nonrestrictive cue only in ambiguous contexts (although this contrast was not directly manipulated in the current study).

4. General discussion and conclusions

The current study aimed at investigating whether *nonrestrictive* contextual information (in the form of modifier number marking) can bias interpretation of ambiguous nouns, and thus disambiguation of NPs, and even an entire syntactic structure.

Firstly, we observed some differences between NP comprehension and sentence production with respect to the ambiguous nouns themselves. When located within a preamble in a sentence production task, a noun ambiguously marked with -ak was more likely to be interpreted as an ergative singular noun, to be used in transitive constructions. This is in line with the previous findings of Yetano et al. (2011), who also found an ergative singular reading was preferred. Although our study

was a sentence production study, while Yetano et al. (2011) focused on sentence comprehension, both studies deal with ambiguously marked phrases within sentences. Processing a sentence means building an entire event, which might trigger the use of transitive constructions and consequently the search for agents. Previous research on the processing of Basque nouns ending in *-ak* has indeed observed a preference for interpreting sentence-initial NPs marked in *-ak* as ergatives. This has been explained by the tendency to ascribe an agentive role to the sentence-initial element (Erdocia et al., 2014; Yetano et al., 2011).

By contrast, when the ambiguous noun is located within an NP in a number judgement task (Experiment 1), where no sentences need be generated, it tends to be interpreted as an absolutive plural, a case used in intransitive constructions. In this type of task, no sentence or event construction is required. Consequently, an interpretation at the level of thematic roles need not be activated so participants resort to the default form. Along these lines, it has been suggested that the absolutive form is the default, citation form for Basque nouns (Hualde & Ortiz de Urbina, 2003), and encountering an isolated NP should directly lead to an absolutive interpretation. Furthermore, the default status of the absolutive form is suggested by evidence that intransitive forms are easier to process across languages (Meltzer-Asscher, Mack, Barbieri, & Thompson, 2015; Thompson & Meltzer-Asscher, 2014), and that the absolutive is acquired before ergative in Basque (Ezeizabarrena, 1996).⁷

We therefore believe that the difference between Experiments 1 and 2 might not be best explained as a difference between comprehension and production, but rather ascribed to the different configurations within which the ambiguous noun was presented. In any case, the results suggest that the interpretation of an ambiguous marking is shaped by task demands and how the system responds to them. Ambiguity resolution should thus be interpreted with respect to the paradigm within which it is being tested.

However, we also demonstrated important similarities between the two paradigms. Namely, general preferences in both configurations seemed to be affected by the number of the modifier. In the

case of ambiguous nouns in a sentential context (Experiment 2), singular modifiers boosted the overall preference for singular nouns; in the case of NP comprehension (Experiment 1), plural modifiers boosted the preference for plural nouns. Furthermore, in both experiments, the overall dispreferred interpretation was boosted by the number of the modifier: the percentage of singular answers numerically increased in the Singular condition of Experiment 1, and the number of absolute plural usages numerically increased in the Plural condition in Experiment 2. These numerical trends are also visible from the Figures 2 and 4. This suggests that both singular and plural modifier markings, although *nonrestrictive*, can indeed act as disambiguating information, regardless of the paradigm.

Moreover, Experiment 2 revealed that both singular and plural modifiers influenced erroneous verb choice, reflecting the well-established phenomenon of number attraction in agreement (e.g. Bock & Miller, 1991). In our case, both singular and plural modifiers acted as *attractors*, although the singular modifier effect was overshadowed by the striking prevalence of plural errors. These errors (mixing *-ak* and *-ek* marked nouns) might be described as a feature of spoken Basque, as we previously mentioned. Hence, the current study shows similar patterns for disambiguation and attraction: the most robust effects were a general preference (in the case of disambiguation) and a common error (in the case of attraction), but the number of the modifier always affects these overall tendencies.

We can therefore conclude that morphological marking on a word in the immediately preceding context strongly affects syntactic disambiguation. This is a novel finding, as we show that syntactic disambiguation is affected by *nonrestrictive* information which doesn't constrain sentence parsing. In contrast to previous studies where relying on nonrestrictive cues generated interpretation difficulties later in the sentence (Bader, 1997; Bader & Meng, 1999), using nonrestrictive information in the current study did not entail processing differences. Indeed, opting for either alternative (e.g. interpreting the ambiguous noun as singular and using the singular transitive construction in the presence of plural modifiers, or vice versa) would result in a perfectly grammatical and semantically congruent sentence.

The important question then is *why* the parser would use *nonrestrictive* information. That is, why would the construction of the entire sentence adhere to the number of the modifier, when not adhering to it has no consequences for processing? To answer this question, it is important to highlight that in the current study design, the head noun contains both number and case ambiguity and that disambiguating this noun is necessary to perform the task (i.e. to provide a number judgement in Experiment 1, and to complete the sentence, in Experiment 2). Importantly, modifier marking is the only accessible information about number. We therefore suggest that the parser uses *any* information that is available to resolve temporary ambiguity. This interpretation is further supported by the lack of any effect of modifier number on unambiguous filler items in both experiments. Although further studies that experimentally contrast ambiguous and unambiguous conditions are necessary to draw any stronger conclusions, our additional analyses suggest that when no disambiguation is necessary, number marking is not used as a disambiguating cue. In their Parallel Function Hypothesis, Bader & Meng (1999) propose that the parser employs certain mechanisms to resolve syntactic ambiguity. We agree with this proposal and suggest that disambiguating mechanisms can be so strong, that they override the features of an ambiguous noun, leading to number attraction errors. This is not so surprising, since in essence, ambiguity is problematic for the linguistic system. The parser has to deal with two possible interpretations, and this increases processing costs. Given reduced processing resources, the parser might be more prone to use *any* available contextual information whatsoever to resolve ambiguity.

Acknowledgements

The authors are grateful to Eneko Antón, Amets Esnal, Andrea Ganchegui, Itzal Uranga, Ian Rion and Manex Lete for their help with stimuli creation, recording and transcription. This research was partially funded by the following grants: P1_2014_1_38 (B.R.) and PI_2016_1_0014 (N.M.) from the Basque Government, PSI2015-65694-P and RTI2018-096311-B-I00 (N.M), as well as RYC-2017-22015 and FFI2016-76432-P_LAMPT (S.M.), funded by the Ministerio de Economía y

Competitividad (MINECO), the Agencia Estatal de Investigación (AEI) and the Fondo Europeo de Desarrollo Regional (FEDER), the Gipuzkoa Fellowship Program (S.M), as well as the Centro de Excelencia Severo Ochoa grant SEV-2015-0490.

Disclosure Statement

The authors report no conflicts of interest.

References

- Acha, J., Laka, I., Landa, J., & Salaburu, P. (2014). EHME: a New Word Database for Research in Basque Language. *The Spanish Journal of Psychology*, *17*, E79.
- Altman, G. T. M., & Steedman, M. (1988). Interaction with context during human speech comprehension. *Cognition*, *30*, 191–238.
- Bader, M. (1997). *Syntactic and Morphological Contributions to Processing Subject-Object Ambiguities*. Unpublished manuscript, University of Jena, Jena, Germany.
- Bader, M., & Meng, M. (1999). *Case attraction phenomena in German*. Unpublished manuscript, University of Jena, Jena, Germany.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing : Keep it maximal. *Journal of Memory and Language*, *44*, 255–278.
- Bates, D., Machler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.
- Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, *23*(1), 45–93.
- Burgess, R. C. (1991). *Interaction of semantic, syntactic and visual factors in syntactic ambiguity resolution* (Unpublished doctoral dissertation). University of Rochester, Rochester, NY, USA.
- de Bruin, A., Carreiras, M., & Duñabeitia, J. A. (2017). The BEST Dataset of Language Proficiency. *Frontiers in Psychology*, *8*, Article 522.
- Eberhard, K. M. (1997). The Marked Effect of Number on Subject–Verb Agreement. *Journal of Memory and Language*, *36*(2), 147–164.
- Erdocia, K., Laka, I., Mestres-Missé, A., & Rodriguez-Fornells, A. (2009). Syntactic complexity and ambiguity resolution in a free word order language: Behavioral and electrophysiological

- evidences from Basque. *Brain and Language*, 109(1), 1–17.
- Erdocia, K., Zawiszewski, A., & Laka, I. (2014). Word Order Processing in a Second Language: From VO to OV. *Journal of Psycholinguistic Research*, 43(6), 815–837.
- Ezeizabarrena, M. J. (1996). *Adquisición de la morfología verbal en euskera y castellano por niños bilingües*. Leioa: Servicio Editorial de la Universidad del País Vasco.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14, 178–210.
- Hualde, J. I., & Ortiz de Urbina, J. (2003). *A Grammar of Basque*. Berlin: Mouton de Gruyter.
- Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition*, 95(1), 95–127.
- Kröger, J. M., Münster, K., & Knoeferle, P. (2017). The Influence of Prosody and Case Marking on Thematic Role Assignment in Ambiguous Action Scenes : Adults versus Children. *CogSci*, 2463–2468.
- Laka, I. (1996). *A Brief Grammar of Euskara, the Basque Language*. University of the Basque Country, Office for the Vicerector for the Basque Language. Retrieved from: <https://www.ehu.es/en/web/eins/basque-grammar>
- MacDonald, M. C. (1994). Probabilistic Constraints and Syntactic Ambiguity Resolution. *Language and Cognitive Processes*, 9(2), 157–201.
- MacDonald, M., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Syntactic ambiguity resolution as lexical ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 123–153). Hillsdale: Lawrence Erlbaum Associates, Inc.
- Malaia, E., Wilbur, R. B., & Weber-Fox, C. (2010). ERP evidence for telicity effects on syntactic

- processing in garden-path sentences. *Brain and Language*, 108(3), 145–158.
- Marslen-Wilson, W. D., & Tyler, L. K. (1987). Against modularity. In J. L. Garfield (Ed.), *Modularity in knowledge representation and natural language understanding* (pp. 37–62). Cambridge, MA: The MIT Press.
- Meltzer-Asscher, A., Mack, J. E., Barbieri, E., & Thompson, C. K. (2015). How the brain processes different dimensions of argument structure complexity: Evidence from fMRI. *Brain and Language*, 142, 65–75.
- Ni, W., & Crain, S. (1990). How to resolve structural ambiguities. In J. Carter (Ed.), *Proceedings of the 20th annual meeting of the North Eastern Linguistic Society* (pp. 414–427). Amherst: University of Massachusetts.
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195-203.
- R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <http://www.R-project.org/>.
- Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension : Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology*, 45(4), 447–481.
- Tabossi, P., Spivey-Knowlton, M. J., McRae, K., & Tanenhaus, M. K. (1994). Semantic effects on syntactic ambiguity resolution: Evidence for a constraint-based resolution process. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance 15: Conscious and nonconscious information processing* (pp. 589–615). Cambridge, MA: The MIT Press.
- Thompson, C. K., & Meltzer-Asscher, A. (2014). Neural mechanisms of verb argument structure processing. In A. Bachrach, I. Roy, & L. Stockall (Eds.), *Structuring the Argument:*

Multidisciplinary research on verb argument structure (Vol. 22, pp. 141–168). Amsterdam:
John Benjamins Publishing Company.

Trueswell, J. C. (1996). The Role of Lexical Frequency in Syntactic Ambiguity Resolution. *Journal of Memory and Language*, 35(4), 566–585.

Trueswell, J. C., & Tanenhaus, M. K. (1991). Tense , Temporal Context and Syntactic Ambiguity Resolution. *Language and Cognitive Processes*, 6(4), 303–338.

Trueswell, J. C., Tanenhaus, M. K., & Garney, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, 33(3), 285-318.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. T. (1993). Verb-Specific Constraints in Sentence Processing : Separating Effects of Lexical Preference From Garden-Paths. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 19(3), 528–553.

Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition*, 99(2), B63-B72.

Yetano, I., Duñabeitia, J., & Laka, I. (2011). *Agent-initial processing preference in Basque: A visual-world eye-movement experiment*. Poster Presented at the 7th International Morphological Processing Conference, San Sebastian, Spain.

Notes

¹ This was confirmed by the current experiment, with only 1.6% of the total absolute use involving objects in Experiment 2.

² see de Bruin et al., 2017

³ In Basque, the uninflected article-less form of the noun is used in this construction (i.e. *mutil*), which conveniently excludes the possibility of the question biasing the answer.

⁴ Adding the frequency of the genitive form (plural/singular) or the NP rating score did not improve model fit, nor did it result in any significant effects.

⁵ To check for a possible effect of Transcriber (the native speaker who transcribed the answers) on our effect of interest, we fitted a model that contained a Condition X Transcriber interaction. The maximally converging model included Condition X Transcriber interaction as a fixed effect, by-participant and by-item random intercept and by-item random slope. There was no interaction between Condition and Transcriber ($z < 0.70, p > 0.40$). Also, adding this variable did not result in a better model fit.

⁶ Adding the frequency of the genitive form (plural/singular), the NP rating score, or the language dominance of the participants did not improve model fit, nor did it result in any significant effects.

⁷ We thank an anonymous reviewer for these observations.

condition	Singular genitive	Plural genitive
	<i>Gaixoaren nebak</i> patient-GEN.SG brother-ERG.SG/ABS.PL ‘the brother(s) of the patient’	<i>Gaixoen nebak</i> patient-GEN.PL brother-ERG.SG/ABS.PL ‘the brother(s) of the patients’

Table 1. An example of an experimental item

	Total	Singular genitive	Plural genitive
Plural answers	73.5%	71%	76%

Table 2. The percentages of plural answers per condition in Experiment 1

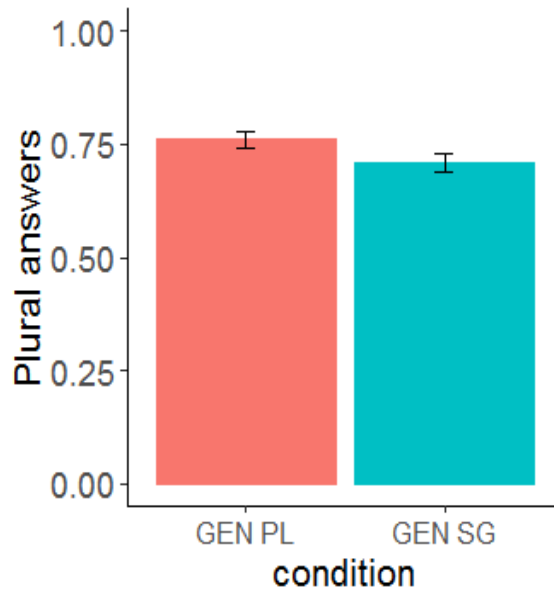


Figure 1. The use of plural answers per condition in Experiment 1. PL GEN=Plural genitive condition, SG GEN=Singular genitive condition

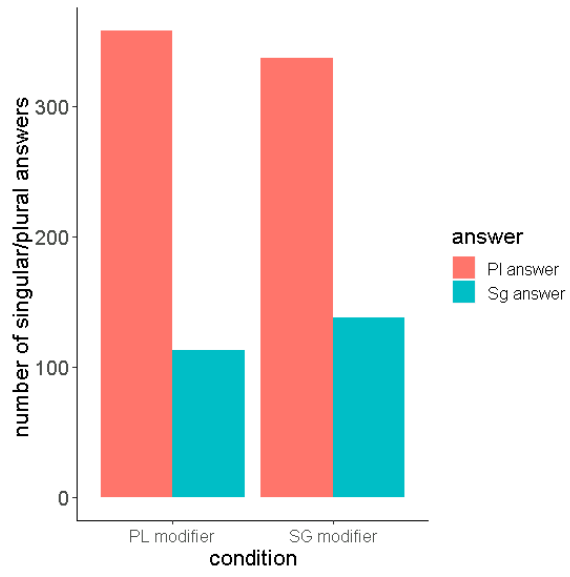


Figure 2. The occurrence of singular and plural answers per condition in Experiment 1. Sg= singular; PI= plural

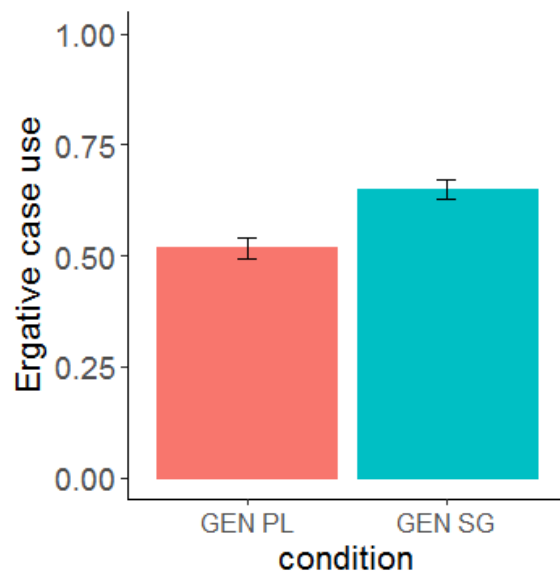


Figure 3. The proportion of the use of ergative singular case per condition in Experiment 2. PL GEN=Plural genitive condition, SG GEN=Singular genitive condition

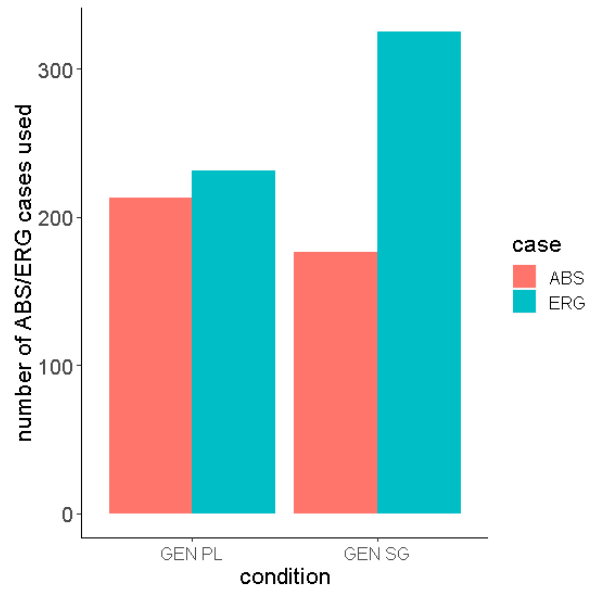


Figure 4. The use of cases per condition in Experiment 2. PL GEN=Plural genitive condition, SG GEN=Singular genitive condition

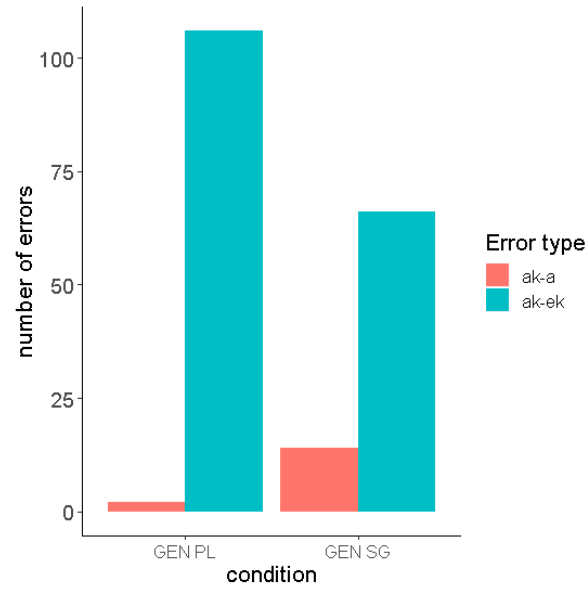


Figure 5. The occurrence of two different types of errors per condition in Experiment 2. PL GEN=Plural genitive condition, SG GEN=Singular genitive condition



Figure 6. The experimental procedure in Experiment 1. The participants were first presented with the ambiguous NPs for 2 seconds in the center of the screen (e.g. *gaixoaren nebak*, “the brother/brothers of the patient”), after which a screen with the question appeared. The question was related to the number of the ambiguous noun (e.g. *Zenbat neba daude?*, “How many brothers are there?”), with the two answer options in the bottom of the screen (*bat*, “one” and *bat baino gehiago*, “more than one”). To answer the questions, the participants used two corresponding keys on the keyboard (the X key for the answer option on the left, and the M key for the answer option on the right)