Pronoun preference of children in a language without typical third-person pronouns

Pronoun preferences of children in a language without typical third-person pronouns

Journal:	*First Language*
Manuscript ID	FLA-16-0004.R2
Manuscript Type:	Original Manuscript
Keywords:	Null pronoun, overt anaphoric form, coreference, topic continuity, topic shift
Abstract:	This study presents comprehension data from 6–7 and 8–10-year-old children as well as adults on the acceptability of null vs. overt anaphoric forms (the demonstrative hura 'that' and the quasipronoun bera '(s)he (him-/herself)') in Basque, a language without true third-person pronouns. In an acceptability judgement task, a developmental change occurred in the preference for hura (Experiment 1): 6–7-year-olds showed a preference for the null pronoun in both topic-shift and topic-continuity contexts, while 8–10-year-olds, like adults, preferred hura in topic-shift contexts and null pronouns in topic-continuity contexts. However, no developmental shift was observed in the preference for bera (Experiment 2): unlike adults, neither 6–7- nor 8–10-year-old children selected bera over null pronouns in topic-shift contexts. They instead showed a general preference for null pronouns, an indication of tolerance for ambiguity – a pattern which differs from prior studies in other null-subject languages where ambiguous pronouns declined with age. The results reveal a different developmental pattern for hura and bera, which may be explained by the more rigid (syntactic) constraints operating on hura in comparison to bera in antecedent choice.

Pronoun preferences of children in a language without typical third-person pronouns

Corresponding author:

Maialen Iraola Azpiroz, The University of Kaiserslautern, Erwin Schrödinger Straße 57, 67633 Kaiserslautern, Germany.

E-mail: maialeniraola@yahoo.es

Maialen Iraola Azpiroz[1,2], Mikel Santesteban[2], Antonella Sorace[3] and Maria-José Ezeizabarrena[2]

[1] Faculty of Social Sciences, The University of Kaiserslautern, Germany

[2] Department of Linguistics and Basque Studies, University of the Basque Country (UPV/EHU), Spain

[3] School of Philosophy, Psychology ad Language Sciences, The University of Edinburgh, UK

Abstract

This study presents comprehension data from 6–7 and 8–10-year-old children as well as adults on the acceptability of null vs. overt anaphoric forms (the demonstrative *hura* 'that' and the quasipronoun *bera* '(s)he (him-/herself)') in Basque, a language without true third-person pronouns. In an acceptability judgement task, a developmental change occurred in the preference for *hura* (Experiment 1): 6–7-year-olds showed a preference for the null pronoun in both topic-shift and topic-continuity contexts, while 8–10-year-olds, like adults, preferred *hura* in topic-shift contexts and null pronouns in topic-continuity contexts. However, no developmental shift was observed in the preference for *bera* (Experiment 2): unlike adults, neither 6–7- nor 8–10-year-old children selected *bera* over null pronouns in topic-shift contexts. They instead showed a general preference for null pronouns, an indication of tolerance for ambiguity – a pattern which differs from prior studies in other null-subject languages where ambiguous pronouns declined with age. The results reveal a different developmental pattern for *hura* and *bera*, which may be explained by the more rigid (syntactic) constraints operating on *hura* in comparison to *bera* in antecedent choice.

Keywords

Null pronoun, overt anaphoric form, coreference, topic continuity and topic shift

Introduction

Pronouns are reference-tracking devices that allow speakers to refer back to already introduced referents (i.e. anaphora) in the discourse or to anticipate referents that will be mentioned later on (i.e. cataphora). The so-called pronominal anaphora is a common linguistic device to avoid repetition of the same linguistic expressions (nominal categories, phrases, etc.). When using third-person pronouns, the speaker signals to the interlocutor mutual familiarity with their referents, since pronouns usually tend to refer to entities that have been previously mentioned in the discourse. Thus, in an ongoing conversation, for effective communication between two interlocutors to happen, the listener must be able to identify the referent of a pronoun, i.e. the antecedent. The identification of the antecedent may be difficult in certain contexts because pronouns are not categorical in interpretation —they do not convey enough referential information on their own but rather their interpretation is conditioned by both linguistic and extra-linguistic information. For this reason, reference assignment presents a challenge to models of natural language processing and also to the development of pronoun-antecedent mappings (O'Grady, 1997).

Early monolingual language acquisition research devoted a great deal of attention to children's mastery of pronouns, e.g. whether children are capable of coordinating knowledge from different domains, since pronoun reference is subject not only to syntactic

but also to contextual constraints. In this regard, dependencies introduced by reflexive pronouns have been reported to be mastered at an earlier stage (e.g. *Mary$_i$ likes herself$_{i/*j}$*) than those introduced by non-reflexive pronouns (*Mary$_i$ likes her$_{j/*i}$*, see Guasti, 2002 for a review), despite the latter being more frequent in child speech than the former (O'Grady, 2005). These differences have been explained in terms of the different principles governing their interpretation; whereas a syntactic dependency is established for the interpretation of reflexives, such a requirement does not hold for non-reflexive pronouns, where coreference with an extrasentential referent involves accessing information beyond syntax (i.e. discourse). Thus, syntactic dependencies relying on the computational language system are more easily acquired than discourse dependencies, and they are less costly in terms of processing, since they are immediately interpretable without requiring pragmatic knowledge (O'Grady, 2005). Syntactic dependencies may also be dependent on individual differences in cognitive control (Sorace, 2011).

More recently, research on the anaphoric dependencies of null-subject languages with bilingual children has also provided evidence for the validity of the "syntax-before-discourse" hypothesis by testing antecedent preferences for null and overt pronouns. In Italian, the overt pronoun, which typically marks a change of referent, is specified for the interpretable feature [+topic shift, +TS] (Sorace, 2000), whereas the null subject usually

signals topic continuity [-topic shift, -TS]. Studies that primarily focused on production data from simultaneous bilingual children have reported a pragmatic deviance consisting of an asymmetric overextension of overt pronouns (the use of an overt pronoun instead of a null pronoun in [-TS] contexts, e.g.: *Mentre Gianni$_i$ mangia lui$_{?i}$ parla al telefono*. 'While Gianni eats he talks on the phone', Sorace & Serratrice, 2009), but not in the other direction, i.e. the use of null pronouns in overt pronoun environments (among others, Serratrice, Sorace, & Paoli, 2004; Hacohen & Schaeffer, 2007). These findings led to the conclusion that bilingual children acquire syntactic conditions for licensing null subjects at an earlier stage than discourse-pragmatic constraints on pronoun realisation. However, studies on comprehension, particularly in bilingual but also monolingual children, have shown a bidirectional non-adult-like extension of the scope of both overt and null pronouns, with the overextension of the null pronoun occurring to a lesser extent (e.g. introduction of a new referent via a null pronoun instead of an overt pronoun in [+TS] contexts, e.g.: *Perché Maria$_i$ é uscita? __$_{?j}$ ha deciso di fare una passeggiata.* 'Why did Maria leave? She decided to go for a walk'; Sorace, 2000).

Both monolingual and bilingual children acquiring a null-subject language have been reported to go through a protracted stage in which they exhibit differential sensitivity to the discourse conditions affecting the selection of appropriate pronominal forms (Shin &

Cairns, 2012 for Spanish; Sorace, Serratrice, Filiaci, & Baldo, 2009 for Italian). However, the pragmatically appropriate use of pronouns in different conditions usually develops over time, with older children (8–10) resembling adult preferences of interpreting null pronouns as coreferent with topical antecedents and overt pronouns as referring to non-topical antecedents. Note, nevertheless, that in some studies younger children and not older ones reproduce adults' pronoun interpretations, e.g. the preference of Greek-speaking 6–7-year-old children for a topical antecedent for null pronouns does not remain stable with increasing age, resulting in a U-shaped development (Papadopoulou, Peristeri, Plemenou, Marinis, & Tsimpli, 2015). Thus, despite considerable research on the interpretation preferences of pronouns in monolingual and bilingual development, there are still a number of open questions regarding the felicitous use of pronominal forms in different discourse contexts. Do school-age children lack syntactic knowledge, discourse knowledge, or both kinds of knowledge, or is it rather the real-time use and updating of the referential mappings in context which results in difficulties (see Sorace, 2011 for a discussion)? Alternatively, is the development of the listener's perspective a crucial requirement for adequate referential choice (Hendricks, Koster, & Hoecks, 2014; Shin & Cairns, 2012)?

In the current study, we examine the different antecedent biases of null and overt pronouns and the developmental trajectory of children as they acquire the discourse

features of null and overt anaphoric subjects by comparing the performance of L1 Basque 6–7-year-olds with that of an older group of children aged 8–10; a control group of adults was used as a baseline to observe the children's developmental pattern towards the target language. The present study aims to investigate the developmental stages in the acquisition of the discourse features linked to null and overt pronouns in Basque, a null-subject language without "true" third-person pronouns in which two overt forms, the demonstrative *hura* 'that' and the so-called quasipronoun *bera* '(s)he (him-/herself)' fulfil the anaphoric functions of personal pronouns in other languages. To that end, we use an acceptability judgement task based on short animations.

Anaphora resolution in languages with two pronominal forms

The literature referred to in the previous section has mainly concentrated on the sensitivity of null-subject language speakers towards the discourse features encoded in null and overt pronominal subjects. However, in certain non-null subject languages, speakers must become familiar with the distribution of two overt pronominal forms in anaphoric use: a personal pronoun and a demonstrative. The pragmatic functions of personal pronouns and demonstratives are closely related because they are used to organise the information flow in the discourse by keeping track of previously mentioned referents. However, in a non-null

subject language such as German, these two pronominal forms display distinct referential properties. Whereas in (1) the personal pronoun *er* is interpreted as referring to the preceding subject antecedent *der Anwalt,* the demonstrative *der* is coreferent with the object antecedent, *Klienten*.

(1) Der Anwalt$_i$ sprach mit einem Klienten$_j$. Da er$_i$/der$_j$ nicht viel Zeit hatte, vereinbarten sie ein weiteres Gespräch nächste Woche.

'The lawyer talked to a client. Since he did not have much time, they agreed to have another meeting next week.'

Example from Diessel (1998, p. 96)

These different patterns of antecedent preferences have been confirmed experimentally in adult native speakers by Wilson (2009) using eye tracking. Whereas the demonstrative made straightforward reference to a postverbal antecedent, the personal pronoun showed ambiguous referential properties. Such divergent behaviour between pronominals was also visible in the time course analysis conducted by Ellert (2013), in which the bias of the demonstrative emerged earlier than that of the personal pronoun (800 vs. 1400 ms after the onset, respectively). The observation that distinct pronominal forms behave differently has also been made in Finnish by Kaiser and Trueswell (2008). According to these authors,

referring expressions can exhibit varying degrees of sensitivity towards a number of different constraints in their preferences for the most appropriate antecedent, as formulated in the Form-Specific Multiple-Constraints approach. Studies on pronouns and demonstratives in Estonian (Kaiser & Vihman, 2006), Dutch (Kaiser, 2011) and English (Brown-Schmidt, Byron, & Tanenhaus, 2005) have also reported similar results.

These differences in the antecedent biases of personal pronouns and demonstratives in non-null subject languages, with the latter exhibiting more definite preferences, parallel the distinct biases observed for null and overt pronouns in null-subject languages (Sorace, 2011). In contrast to the overwhelming preference of the null pronoun for a topical antecedent (see, among others, Alonso-Ovalle et al., 2002 for Spanish, Carminati, 2002 for Italian and Mayol, 2009 for Catalan), the overt pronoun's bias towards a non-topical antecedent is not always uniform. The variability in the overt pronoun's resolution preferences depends, for example, on the number of referents in the sentence (Carminati, 2002), the language under study (Italian vs. Spanish, see Filiaci, Sorace, & Carreiras, 2014), and whether the sentence shows anaphoric or cataphoric dependencies (Kraš, Sturt, & Sorace, 2014).

Development trajectories in the acquisition of the discourse features of pronouns

In recent years, a substantial number of studies have been conducted on monolingual and bilingual children's developmental steps towards sensitivity to distinct discourse conditions influencing the selection of pronominal forms. In an acceptability judgement task conducted by Sorace et al. (2009) in Italian, overt pronouns were significantly more often accepted in [-TS] contexts by 6–7-year-old Italian monolingual children and by both Italian-English and Italian-Spanish bilingual children, in contrast to both 8–10-year-old monolingual child and adult controls. To a lesser extent, some pragmatically inappropriate null subject pronouns were also selected in [+TS] contexts by bilinguals regardless of age and language combination and less often by 6–7-year-old monolingual children. Such data suggest that Italian monolingual (and also bilingual) children learn to avoid ambiguity (avoidance of null pronouns in [+TS] contexts) earlier than redundancy (avoidance of redundant overt pronouns in [-TS] contexts).

Redundancy persisting in the course of sensitivity towards discourse conditions on subject pronouns has also been attested in Spanish. Shin and Cairns (2012) obtained preferences for null and overt pronouns in short stories presented to Mexican-Spanish school-aged children. Their results indicated that while 6–7-year-old children did not show

a preference for overt over null pronouns in [+TS] contexts, children from age 8 onwards resembled adults in showing a preference for overt pronouns in [+TS] contexts. However, no preferences for the null (vs. overt) pronoun in [-TS] contexts were exhibited by any child group, not even the group of 14–15-year-olds, who still accepted redundant overt pronouns as referring back to topical antecedents.

The acceptability status of overt (*hura* or *bera*) and null subject pronouns in [+TS] and [-TS] contexts has also been tested in Basque (Iraola Azpiroz, Santesteban & Ezeizabarrena, 2014, Iraola Azpiroz, 2015) by adapting Sorace et al.'s (2009) materials. In contrast to prior studies in Italian and Spanish in which redundant pronouns posed more difficulties for children than ambiguous pronouns, 6–7-year-old L1 Basque-speaking children showed more tolerance for ambiguity by accepting infelicitous null pronouns (the always grammatical option) in [+TS] contexts. This was more evident when the null pronoun was contrasted (in a two-choice preference task) against *bera* than when it was contrasted against *hura.* In addition, preferences for null vs. *bera* were not affected by discourse context, whereas preferences for null vs. *hura* definitely were. This finding suggests that the mastery of the discourse features of *hura* and *bera* might involve different acquisition patterns.

Third-person reference in Basque

A characteristic which distinguishes Basque from other null-subject languages previously studied, and particularly important for the phenomenon under study, is that Basque lacks "true" third-person pronouns. The pronominal inventory only consists of the first and the second person, and therefore Basque has been regarded as a *two-person language* (Bhat, 2004). A null pronoun is the most frequent option in Basque for referring to a third party, except for focused contexts or when a new topic is introduced. In the latter cases, the distal demonstrative *hura* 'that' is employed in Basque. In addition to *hura*, another overt pronominal form, namely the quasipronoun *bera* '(s)he (him-/herself)', acts as a third-person pronoun and competes with the pronominal uses of *hura* (de Rijk, 2008). According to de Rijk, the quasipronoun status of *bera* stems from presence of the root *ber-* (with the adjectival meaning of 'the same'. Despite some overlap in their distributions, *hura* and *bera* show different behaviour in antecedent preferences in intrasentential anaphora contexts, as in (2).

(2) a. Peruren$_i$ amak bera $_{i/?j}$ ikusi du.

'Peter$_i$'s mother has seen him $_{i/?j}$.'

b. Peruren$_i$ amak hura$_{*i/j}$ ikusi du.

'Peter$_i$'s mother has seen him$_{*i/j}$.'

Examples from Eguzkitza (1986, p. 31)

Bera usually refers to persons or objects previously mentioned in the discourse, like *Peter* in (2a), despite not discarding the possibility of referring to a third party. In contrast, *hura* cannot make reference to an intrasentential antecedent (Laka, 1996), hence it refers to an extrasentential referent in (2b).

Pronoun resolution preferences of children in Basque

In order to advance our understanding of the development of sensitivity to discourse constraints on the use of pronouns in null-subject languages, the current study investigates Basque children's preferences for null and overt subject pronouns in [+TS] and [-TS] contexts, and compares them to those of adults. The participants' interpretations of two overt forms (the demonstrative *hura* 'that' and the quasipronoun *bera* '(s)he (him-/herself)') in contrast to the null pronoun are analysed in an acceptability judgement task based on short animations originally designed by Sorace et al (2009). In contrast to the

crosslinguistic developmental pattern in the sensitivity of the discourse features attached to pronouns whereby pragmatically inappropriate null pronouns decrease earlier than infelicitous overt pronouns, previous data from Basque have suggested that children have greater difficulty in avoiding ambiguous null pronouns (Iraola Azpiroz, Santesteban & Ezeizabarrena, 2014, Iraola Azpiroz 2015). Thus, the developmental pattern described in the literature thus far may not be generalisable to Basque. The current paper extends the samples of previous studies on children's pronoun interpretation in Basque in two directions: a) in the number of participants in the 6–7-year-old group (in order to test the consistency of the pattern previously found) and b) in age groups, by including data from an older cohort (8–10-year-olds).

The aims of the current study are twofold. First, it investigates whether Basque-speaking children's antecedent preferences for choices of the two overt anaphoric forms *hura* and *bera* differ from those of null pronouns, consistent with the patterns observed crosslinguistically. More importantly, this study seeks to reveal the developmental paths towards adult-like preferences for null and overt pronouns between ages 6–7 and 8–10. The first prediction is that *hura* and *bera* will be preferred in [+TS] contexts by older children and adults in accordance with the preferences shown by native speakers of null-subject languages, whereas null pronouns will be preferred in [-TS] contexts. Based on our

previous data from 6–7-year-olds, a second prediction states that Basque-speaking children will have more difficulty in discarding null pronouns in contexts involving a switch of reference than in discarding overt pronouns in contexts of reference maintenance. Finally, in accordance with our prior results reporting different discourse context effects for *hura* and *bera*, our third prediction is that children will reach adult standards at an earlier stage for the distribution of the demonstrative *hura* than for that of the quasipronoun *bera.*

Experiment 1: Preferences for *hura* 'that' vs. null subject pronouns

Participants

A group of 38 children with Basque as L1 (age range 6;3-7;4, mean 6;8, 22 F and 16 M), of whom 19 previously participated in the study by Iraola Azpiroz (2015: Experiment 5) and Iraola Azpiroz, Santesteban & Ezeizabarrena (2014: Experiment 2; N=19), took part in the experiment. In addition, a group of 26 older children (age range 8;5-10;4, mean 9;5, 10 F and 16 M) participated. The control group consisted of fourteen Basque native adults (age range 17;2-57;5, mean 28;3). The data of one child from the younger group of children was excluded from the analysis because he responded incorrectly to more than 50% of the filler items. This threshold was implemented to ensure that participants understood the aim of the task. The children were being raised in Basque-speaking families living in Tolosa, a town located in a Basque-dominant sociolinguistic environment in the Spanish province of

Gipuzkoa within the Basque Autonomous Community with an average rate of 50-75% of Basque knowledge (Altuna, 2007). The adult participants came from the same town as the children. Both the children and the adults were native speakers of Basque and used Basque predominantly in their family and social environments, but they also had daily exposure to Spanish.

Materials

The materials and methodology used by Sorace et al. (2009) for Italian were adapted to Basque. Experimental items consisted of short video clips with four characters (Mickey Mouse, Minnie Mouse, Donald Duck and Daisy) with whom the children were familiar. The videos showed one of the two characters in the foreground performing a one-referent action. The action was commented upon, either by the character involved in the action ([-TS] context, see (3)) or by a second character who saw the action ([+TS] context, see (4)). The action was followed by each of the two characters in the background stating what had occurred using either a null (3a, 4a) or an overt anaphoric form (*hura* 'that') (3b, 4b). The null pronoun was expected to be chosen in [-TS] contexts and the overt anaphoric form in [+TS] contexts.

(3) [–TS] context

(Minnie and Daisy in the foreground; Mickey and Donald in the background)

Minnie falls and says: *Erori egin naiz!*

'I've fallen down!'

a. Donald: *Minniek ___erori dela esan du.*

'Minnie said that ___ has fallen down.'

b. Mickey: *Minniek hura erori dela esan du.*

'Minnie said that she has fallen down.'

(4) [+TS] context

(Minnie and Daisy in the foreground; Mickey and Donald in the background)

Daisy falls and Minnie says: *Daisy erori egin da!*

'Daisy has fallen down!'

a. Donald: *Minniek ___ erori dela esan du.*

'Minnie said that ___ has fallen down.'

b. Mickey: *Minniek hura erori dela esan du.*

'Minnie said that she has fallen down.'

The task included a total of 16 experimental items (8 items per [-TS] and [+TS] condition) and 10 filler items.

Procedure

Child participants were individually tested at school in a quiet room. The materials were presented in a PowerPoint presentation on a laptop, and the children's responses were recorded on an answer sheet. The participants were told that the four characters had started to learn Basque, but that they sometimes made mistakes. Note that most children were familiar with these characters from watching cartoons in Spanish on TV. The experimenter piqued the child's interest by saying that she knew that the child spoke Basque very well, and that she needed his or her help to decide which character located in the background spoke "better" Basque. Participants indicated their preferences by pointing at one of the characters. Trials were pseudorandomised, and care was taken to ensure that no more than two trials of the same experimental condition were presented consecutively.

Data analysis

A mixed-effects logistic regression model was created with pronoun choice (preference for null or overt forms) as the dependent variable, and with discourse context ([–TS] vs. [+TS]) and group (younger children vs. older children vs. adults) as fixed-effect variables

(treatment coded and centred, with [-TS] and the younger group as baseline intercepts). The maximal random effect structures without convergence problems (justified by χ^2-test model comparisons) are reported, which in all cases resulted in models without any by-participant or by-item random slopes. Because treatment coding compares the baseline intercept (e.g. the younger group) with the other conditions of the variable (vs. older children and adults), the performance of older children vs. adults was compared by running the best fit model with older children as the intercept. Finally, whenever context by group interactions are reported, simple [-TS] and [+TS] context models including group as the only fixed effect were created in order to determine whether the group effects were significant for these two context levels.

Results

Table 1 reports the raw data and mean percentages of selection of *hura* and null pronouns in [±TS] contexts.

[Insert Table 1.]

The main effect of context revealed that younger children had a stronger preference for the use of *hura* in [+TS] than in [–TS] contexts (see Table 2), a preference also displayed by both older children (β = 3.842, SE = .362; z = 10.588, p < .001) and adults (β

= 4.720, SE = .484; z = 9.738, p < .001), as shown by releveled models. The non-significant main effect of "Group: older children" in Table 2 reveals that younger and older children had a similar overall preference to select *hura* vs. null pronouns. However, the significant main effects of group that compared the performance of younger children vs. adults ("Group: adults" effect in Table 2) and older children vs. adults (releveled model comparison: β = .950, SE = .438; z = 2.167, p = .030) revealed that both younger and older children groups had a weaker tendency to select *hura* than adults. Finally, the significant interactions between context and group (younger vs. older and younger vs. adults; see Table 2) revealed that both older children and adults showed a greater context effect (stronger preference for *hura* in [+TS] contexts) than younger children. However, the non-significant context by group interaction of the releveled model indicated similar context effects for older children and adults (β = .879, SE = .583, z = 1.506, p = .132).

With the aim of exploring the context by group significant interactions, the [-TS] simple models determined that in this context, younger children exhibited a stronger preference to select *hura* than older children (β = -1.867, SE = .623, z = -2.995, p = .002) and adults (β =-1.894, SE = .807, z = -2.348, p = .018), whereas no differences were found between older children and adults (β = .-.027, SE = .869, z = -.032, p = .975). In contrast, in the [+TS] context, younger children displayed a weaker preference for *hura* than older

children and adults (β = 1.89, SE = .492, z = 3.854, $p < .001$ and β = 4.041, SE = .752, z = 5.379, $p < .001$, respectively), as did older children in comparison to adults (β = 2.15, SE = .756, z = -.032, $p < .01$).

The results thus showed that there was a developmental progression towards the adult-like preference for *hura* in [+TS] contexts. In Experiment 2, we investigated whether a similar developmental pattern would also be observed for the overt anaphoric form *bera*, whose scope seems to be larger than that of *hura* according to descriptive grammars.

Experiment 2: Preferences for *bera* '(s)he (him-/herself)' vs. null subject pronouns

Participants

Forty-four young Basque native children (age range 6;1-7;4, mean 6;6, 26 F and 18 M) participated in the experiment. This sample of children aged 6–7 was an extension of the sample in Iraola Azpiroz (2015: Experiment 6; N =23). An older group of 27 children (age range 8;5-10;4; mean 9;2, 11 F and 17 M) also took part in the study. The control group consisted of ten Basque native adults (age range 18;0-20;0, mean 19;0). All participants were different from those in Experiment 1.

Materials and Procedure

The materials and procedure were identical to those used in Experiment 1, except that in the overt pronoun condition, *bera* was used instead of *hura*.

Results

The raw data and mean percentages of selection of *bera* and null pronouns in [±TS] contexts are provided in Table 1. The main effect of context was non-significant neither for young children (see Table 2) nor for older children (releveled model with older children as baseline intercept: β = 241, SE = .339; z = .710, p = .477), which indicates that they showed no preference for the use of *bera* either in [+TS] or [–TS] contexts. In contrast, the significant main effect of context observed by adults indicated that they had a stronger preference for the use of *bera* in [+TS] than in [–TS] contexts (releveled model with adults as baseline intercept: β = 2.977, SE = .494; z = 6.020, p < .001). The main effect of group comparing younger children vs. adults was non-significant suggesting that younger children and adults exhibited a similar overall preference to select *bera*. However the main effects of group comparing older children vs. younger children (see Table 2) and older children vs. adults (releveled model comparison: β = .709, SE = .338; z = 2.099, p = .035) resulted significant with older children selecting *bera* less often than both younger children and

adults. Finally, the significant interactions between group and context revealed different (reversed) context effects for young children compared to both older children and adults (see Table 2). In the case of younger vs. older children, despite both children's groups showing non-significant clear preferences, whereas younger children showed a tendency to prefer *bera* in [-TS] contexts, older children showed a tendency to prefer *bera* in [+TS] contexts. Similarly, in the case of younger children vs. adults, in contrast to the non-preference of younger children, adults showed a clear preference to select *bera* in [+TS] contexts. In addition, the significant group by context interaction of the releveled model indicated that the preference for *bera* in [+TS] contexts was stronger for adults than for older children ($\beta = 2.736$, SE = .481, $z = 5.679$, $p < .001$).

In the analysis of the group by context interactions, the [-TS] simple models determined that in this context, younger children exhibited a stronger preference to select *bera* than both older children ($\beta = -1.069$, SE = .351, $z = -3.046$, $p = .002$) and adults ($\beta = -1.794$, SE = .537, $z = -3.340$, $p < .001$), whereas no differences were observed between older children and adults ($\beta = -.725$, SE = .565, $z = -1.283$, $p = .199$). In contrast, in the [+TS] context, younger and older children displayed similar preferences for *bera* ($\beta = -.418$, SE = .330, $z = -1.264$, $p = .206$), which in both cases was weaker than that of adults ($\beta = 1.91$, SE = .484, $z = 3.957$, $p < .001$ and $\beta = 2.33$, SE = .516, $z = 4.517$, $p < .001$,

respectively).

Discussion

Previous research on the discourse features affecting subject pronouns in Italian and Spanish (Shin & Cairns, 2012, Sorace et al. 2009) have revealed a developmental trajectory towards adult-like preferences. In the present study, we examined the pronominal form (null vs. overt) preferences of a group of Basque native adults and two groups of 6–7- and 8–10-year-old Basque native children in [-TS] and [+TS] contexts by means of an acceptability judgement task in order to identify any developmental progression. The first prediction that null and overt pronouns would show distinct biases in the older group, similar to what has been observed in other null-subject languages, has been borne out in adults (in both experiments), but only partially in the older child group (only in Experiment 1), and not at all in the younger child group. The second prediction that children would demonstrate more difficulty with ambiguity than redundancy has turned out to be accurate. Finally, the third prediction that there would be more significant developmental changes in the resolution of the demonstrative *hura* from younger (6–7) to older (8–10) groups of children than in that of the quasipronoun *bera* has also been confirmed.

Adult preferences

Adults preferred null pronouns in the [–TS] condition but overwhelmingly selected the demonstrative *hura* in the [–TS] condition. For *bera*, adults also showed a preference for a [+TS] interpretation, although this preference was weaker than the one showed for *hura,* suggesting that the two pronominal forms do not have identical referential dependencies.

The differences in the interpretation of distinct pronominal forms such as personal pronouns and demonstratives have been addressed by Cardinaletti and Starke (1999) in terms of Binding Theory (Chomsky, 1981). According to Principle B, pronouns cannot take as antecedent a referent in their governing category. In (4b) *Minnie said that she has fallen down*, Principle B does not forbid pronouns in complex sentences from selecting an antecedent outside of their finite clause. Thus, coreference with the preceding topical antecedent is possible for personal pronouns and hence for the quasipronoun *bera*. On the other hand, in line with the crosslinguistic tendency of demonstratives to take a non-topical antecedent (Diessel, 1999), coreference between the preceding topical antecedent and *hura* is not allowed. Falling under Principle C, *hura* behaves like a lexical item and must therefore be free from any c-commanding antecedent (Eguzkitza, 1986). Adults' stronger preferences for the contextual appropriateness bias of *hura* than for that of *bera* mirror the response patterns found in languages with two anaphoric forms (Ellert, 2013; Wilson, 2009)

and support the predictions of the Form-Specific Multiple-Constraints approach (Kaiser & Trueswell, 2008). At the same time, experimental data support the disjoint interpretation of *hura* mentioned in descriptive grammars of Basque — not allowing an intrasentential antecedent — and the more flexible reading of the quasipronoun *bera*, which does not categorically disallow coreference with a topical antecedent.

Child preferences

In both younger and older children's groups, discourse context only had an effect in the contrast between the demonstrative *hura* and the null pronoun (Experiment 1). Whilst the null pronoun was preferred in the [–TS] condition by both children's groups, there was a developmental trend in the degree of sensitivity towards the [+TS] feature attached to *hura*: although the 8–10-year-olds still did not select *hura* in [+TS] contexts to the same degree as adults did, they showed a significantly stronger preference than 6–7-year-olds. On the other hand, the choice between *bera* and null pronouns (Experiment 2) was not dependent on the mapping between a particular pronominal form and the felicitous pragmatic considerations for contextual appropriateness, with both younger and older groups showing a preference for the null pronoun in both contexts. Thus, the two overt anaphoric forms apparently have different developmental patterns. Furthermore, contrary to prior studies in which Italian and Spanish children seemed to have more difficulties when dealing with redundancy (avoiding

inappropriate overt pronouns in [-TS] contexts) than ambiguity (avoiding infelicitous null pronouns in [+TS] contexts, Shin & Cairns, 2012; Sorace et al., 2009), Basque children appear to be more delayed in learning to avoid ambiguity, as shown in Experiment 2.

One reason for children to exhibit a more protracted developmental course for a particular pronoun than for another could be that the former is more infrequent in the input in comparison to the latter. The pervasive nature of frequency effects across domains in children's first language acquisition has been discussed by Ambridge, Kidd, Rowland and Theakston (2015), the idea being that the most frequent forms are learned first. Following the same line of reasoning, based on her analysis of the different patterns constraining Spanish subject expression, Shin (2015) argues that because subject pronoun omission is more frequent than expression, children do not receive enough positive evidence from the input to determine how and when to use pronouns first. She concludes that children may require a long time to fully acquire the correct usage of input-driven structured variation such as subject pronouns. However, she points out that Mexican-Spanish children receive plenty of evidence with regard to where [+TS] contexts and not [-TS] contexts favour pronouns, and such children are already sensitive to this constraint by age 6–7. For our results, frequency effects may explain why Basque 6–7-year-olds' learner default (i.e. the preferred option) is the null pronoun, mostly used in [-TS] but also in [+TS] contexts. In

fact, multiple (subject and object) argument drop is a frequently employed option in Basque. However, we believe that input frequency is not the most likely determining factor for the non-developmental trajectory of *bera* as opposed to *hura*, in particular because, according to the largest written corpus available in Basque (EHME, Acha, Laka, Landa, & Salaburu, 2014), *hura* and *bera* have similar frequency values (1124 and 1104 appearances per million words, respectively). These two forms might also be relatively infrequent in the children's input, as suggested by a small corpus of 1000 sentences extracted from a dialogue between a 2-year-old child and his father, in which not a single example of *hura/ha(re)k* 'that-absolutive/ergative' was reported in anaphoric use, and only one example of *bera* in the adult speech (Ezeizabarrena, 2009). Thus, although these observations need to be considered with caution, we suggest that input frequency might not be a determining factor for the different developmental trajectories of *bera* and *hura*.

A more convincing explanation for the difference between children's resolution of *hura* and *bera* in relation to adult pronoun resolution strategies is the different nature of the anaphoric dependency in which both overt pronouns are involved. Adults' and older children's non-selection of *hura* in [–TS] contexts is motivated by the impossibility of coreference between *hura* and the c-commanding subject antecedent, which in turn results in resolution of the dependency at the first opportunity (Efficiency Requirement, O'Grady,

2005). In contrast, the correct interpretation of *bera* requires a longer search for an appropriate antecedent, since Principle B applied to personal pronouns does not offer guidance with only discourse constraints being available. The results obtained thus suggest that less costly anaphoric dependencies such as the syntactic dependency between *hura* and the subject antecedent are acquired earlier by children than discourse dependencies with no immediate resolution, as is the case for *bera*. Thus, the current findings support the previously attested pattern that syntax comes before discourse in development, providing further evidence for the widely discussed Interface Hypothesis (Sorace & Filiaci, 2006; see also Sorace, 2011 for discussion).

Another possibility for the differences observed in the developmental patterns of *hura* and *bera* is the more stable behaviour of *hura* vs. the ambiguity of *bera* in the adult language. *Hura* shows a stronger tendency to choose a referent more uniquely (*rigidity*) by more efficiently limiting the number of potential antecedents than *bera*, which has several readings: coreference with a proximate referent in the discourse — overlapping with the referential properties of the null pronoun — or with a third party. In addition, the disjoint interpretation of *hura* may be more salient for children because demonstratives are among the first words children learn (always among the first fifty, and usually among the first ten words (Clark, 1978)), and children are aware that demonstratives usually (as deictics) make

reference to entities in local contexts. The early production of demonstratives in child language is driven by their communicative function for establishing joint attention — a prerequisite for communication and language — together with depicting pointing (Diessel, 2012). The pattern of acquiring the more restricted form earlier than the less restricted one has also been observed in the acquisition of the anaphoric use of German demonstratives and personal pronouns (Bittner & Kuehnast, 2012). 8–10-year-olds' general preference for null pronouns may have been affected by the difficulties in restricting the scope of potential antecedents for *bera*.

One more factor that should not be disregarded in this multifaceted context is children's difficulty in taking another person's perspective, i.e. children were insensitive to the shift of topic marked by the speaker in the reported speech context of the items tested. In fact, there is evidence that referential functions such as *switching* (introducing a new character) and *maintenance* (referring to the same character) are not equally challenging for children. Canadian English-speaking 7–8-year-old children, for example, exhibited lower adequacy levels in switching than in maintenance functions in narrative practice (Colozzo & Whitely, 2014). The ability to take perspective requires a long time for children to learn, as has been observed in several studies on the interpretation of pronouns, and only develops with increasing linguistic experience and cognitive capacity (Hendricks et al., 2014; Shin &

Cairns, 2012). However, if this factor affected children's interpretation of the subject pronoun, the different developmental patterns found for *hura* and for *bera* would still remain unexplained. Further research is needed to determine the age at which children show adult-like preferences for *bera* in [+TS] contexts. However, it is very likely that children do not attain adult-like control over the distribution of *bera* before age 14–15, as suggested by several studies analysing children's systematic use of referring expressions (Karmiloff-Smith, 1986; Hickmann, 2003; Shin & Cairns, 2012). The present study is based on children's comprehension of the discourse features attached to pronouns, but because several studies have argued that pronoun production precedes pronoun comprehension (see Hendricks, 2014 for a review), future studies with production data are needed to shed light on whether Basque-speaking children exhibit the correct production of *bera* and *hura* at an earlier stage than what present comprehension data suggest. This would be very valuable to further support the existence of an asymmetry between children's production and comprehension of pronouns.

Conclusion

In an acceptability judgement task, the general preference of Basque-speaking 6–7-year-old children for the null pronoun regardless of discourse context indicated that the interpretable

[+TS] feature, which is mapped onto *hura* and to a lesser extent onto *bera* in the adult grammar, is still underspecified at that age. The responses of the group of older children revealed a developmental change in the resolution of *hura*: like adults, these 8–10-year-old children showed,different antecedent choice preferences between null and overt pronouns in Basque, in line with Carminati (2002). In contrast, children's resolution of *bera* did not change significantly over the age ranges studied, with 8–10-year-olds still not showing a clear interpretative preference. Thus, sensitivity to the discourse conditions of the two overt anaphoric forms seems to follow two different trajectories, with the [+TS] feature of the demonstrative *hura* 'that' emerging earlier than that of the quasipronoun *bera* '(s)he him/herself'. *Hura,* being constrained by syntactic principles, conveniently reduces the number of possible antecedents, as shown in adults' more robust resolution preferences compared to the more flexible interpretation of *bera.* Differences in the frequency of the two overt anaphoric forms and null pronouns appears to be compatible with an earlier emergence of the discourse constraints attached to the null pronoun, but frequency effects for the acquisition of the contextual conditions affecting the distribution of *hura* and *bera* require more empirical support.

Acknowledgements

We are very grateful to the pupils and teachers of the Basque schools *Laskorain ikastola* and *Uzturpe ikastola*, as well as to Kepa Erdozia, Oxel Uribe-Etxeberria, Amaia Munarriz and Mikel Iraola for their helpful contributions. We also wish to thank the editor and the reviewers for their insightful comments.

Funding

This work was supported by the European Union's Seventh Framework Programme for research, technological development and demonstration (Grant No. 613465); the Basque Government (IT665-13, IT-676-13, POS2015_2_0008); the University of the Basque Country (UFI11/06) and the Ministries of Science and Innovation & Economy and Competitiveness (FFI2012-37884-C03-02, FFI2014-55733-P, FFI2015-68589-C2-1-P (MINECO/FEDER), RyC-2013-14722).

Table 1. Raw data, percentages (in brackets) and *SD*s (in italics) per cell for overt pronouns, null pronouns and miscellaneous responses per group in each experimental condition of Experiments 1 and 2.

	[-TS] contexts			[+TS] contexts			Context effect[1] ([+TS] minus [-TS])
	Overt	Null	Misc	Overt	Null	Misc	Overt
Experiment 1: *HURA*							
6–7-year-olds	75 (24.7%) *43.18%*	229 (75.3%)	0	110 (36.5%) *48.23%*	191 (63.5%)	3	35 (11.8%)
8–10-year-olds	15 (7.2%) *25.93%*	193 (92.8%)	0	138 (67.6%) *46.89%*	66 (32.4%)	4	123 (60.4%)
Adults	11 (9.8%) *29.89%*	101 (90.2%)	0	97 (89.0%) *31.44%*	12 (11.0%)	3	86 (79.2%)
Experiment 2: BERA							
6–7-year-olds	154 (43.8%) *49.67%*	198 (56.3%)	0	126 (36.2%) *48.12%*	222 (63.8%)	4	-28 (-7.6%)
8–10-year-olds	55 (25.7%) *43.80%*	159 (74.3%)	2	63 (29.3%) *45.62%*	152 (70.7%)	1	8 (3.6%)
Adults	13 (16.3%) *37.12%*	67 (83.8%)	0	58 (72.5%) *44.93%*	22 (27.5%)	0	45 (56.2%)

[1]The *Context effect* column shows the participants' preference for overt pronouns, with positive values indicating a preference for overt pronouns in [+TS] contexts and negative values indicating a preference for overt pronouns in [-TS] contexts.

Table 2. Generalized linear mixed models for Experiments 1 and 2.

	ß	*SE*	*z*	*p*
Experiment 1 (HURA)				
(Intercept)	-1.037	.204	-5.082	< .001
Context: [+TS]	.718	.222	3.235	.001
Group: older children	.010	.325	.033	.973
Group: adults	.961	.406	2.367	.017
Interaction: [+TS]/older	3.123	.402	7.770	< .001
Interaction: [+TS]/adults	4.002	.515	7.761	< .001
Experiment 2 (BERA)				
(Intercept)	-.469	.179	-2.611	.009
Context: [+TS]	-.331	.300	-1.104	.269
Group: older children	-.646	.213	-3.026	.002
Group: adults	.062	.318	.197	.844
Interaction: [+TS]/older	.573	.281	2.036	.041
Interaction: [+TS]/adults	3.309	.457	7.238	<.001

Reference

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.Alonso-Ovalle, L., Fernandez-Solera, S., Frazier, L., & Clifton, C. (2002). Null versus overt pronouns and the topic-focus articulation in Spanish. *Rivista Linguistica/ Italian Journal of Linguistics* 14: 151-170.

Altuna, O. (2007), Euskal Herriko kale erabileraren neurketa.Ikerketaren nondik norakoak. *BAT aldizkaria*, 64, 21-31.

Ambridge, B., Kidd, E., Rowland, C.F. & Theakson, A.L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language* 42: 239-273.

Bhat, D. N. S. (2004). *Pronouns*. Oxford: Oxford University Press.

Bini, M. (1993). La adquisición del italiano: más allá de las propiedades sintácticas del parámetro pro-drop. In J. Liceras (ed.), *La lingüística y el análisis de los sistemas no nativos*, pp. 126-139. Ottawa: Dovehouse.

Bittner, D., & Kuehnast, M. (2012). Comprehension of intersentential pronouns in child German and child Bulgarian. *First Language* 32: 176-204.

Cardinaletti, A., & Starke, M. (1999). Responses and demonstratives. In H. van Riemsdijk (ed.) *Clitics in the languages of Europe*, pp. 280-289. Berlin/New York: Mouton de Gruyter.

Carminati, M. N. (2002). *The processing of Italian subject pronouns*. Unpublished doctoral dissertation. University of Massachusetts at Amherst, USA.

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

Clark, E. V. (1978). From gesture to word: On the natural history of deixis in language acquisition. In J. S. Bruner & A. Garton (eds.), *Human growth and development: Wolfson College lectures 1976*, pp. 85-120. Oxford: Oxford University Press.

Colozzo, P., & Whitely, C. (2014). Keeping track of characters: Factors affecting referential adequacy in children's narratives. *First Language* 3: 155-177.

Diessel, H. (1999). *Demonstratives: Form, function, and grammaticalization*. Amsterdam: John Benjamins.

Diessel, H. (2012). Deixis and demonstratives. In C. Maienborn, K. von Heusinger, P. Portner (eds.), *An International Handbook of Natural Language Meaning*, Vol. 3, 2407-2431. Berlin: Mouton de Gruyter.

Eguzkitza, A. (1986). *Topics in the syntax of Basque and Romance*. Doctoral dissertation. UCLA, USA.

Ezeizabarrena, M.J. (2009). Development in language mixing: early Basque-Spanish bilingualism. In J.Grinstead (ed), *Hispanic child languages: typical and impaired development*, pp. 57-89.Amsterdam: Benjamins.

Filiaci, F., Sorace, A., & Carreiras, M. (2014). Anaphoric biases of null and overt subjects in Italian and Spanish: a cross-linguistic comparison. *Language and Cognitive Processes* 29: 825-843.

García Alcaraz, E., & Bel, A. (2015). The acquisition of co-referential properties of pronouns in bilingual and L2 Spanish speakers. *Proceedings of the 39th Boston University on Language Development. Online Proceedings Supplement*. Retrieved from: http://www.bu.edu/bucld/files/2015/06/Garcia-Alcaraz.pdf

Guasti, M.T. (2002). *Language acquisition: The growth of grammar*. Cambridge, MA: MIT Press.

Hacohen A., & Schaeffer, J. (2007). Subject realization in early Hebrew/English bilingual acquisition: The role of crosslinguistic influence. *Bilingualism: Language and Cognition* 10: 333-344.

Hendriks, P. (2014). *Asymmetries between language production and comprehension* (Studies in Theoretical Psycholinguistics 42). Heidelberg: Springer.

Hendricks, P., Koster, C., & Hoecks, J. (2014). Referential choice across the lifespan: Why children and elderly adults produce ambiguous pronouns. *Language, Cognition and Neuroscience* 29: 391-407.

Hickmann, M. (2003). *Children's Discourse; Person, Space and Time across Languages*. Cambridge: Cambridge University Press.

Iraola Azpiroz, M., Santesteban, M. & Ezeizabarrena, M.J. (2014). The discursive feature specification of the pronoun *hura* in child (L1 and L2) and adult Basque. *Selected Proceedings of the Romance Turn V. Workshop on the Acquisition of Romance Languages*, pp. 134-157. Newcastle: Cambridge Scholars Publishing.

Iraola Azpiroz, M. (2015). *Anaphora resolution in children and adults: An experimental study of mature speakers and learners of Basque*. Tübingen: Günther Narr-Verlag.

Kaiser, E. (2011). Salience and contrast effects in reference resolution: The interpretation of Dutch pronouns and demonstratives. *Language and Cognitive Processes* 26: 1587-1624.

Kaiser, E., & Trueswell, J. C. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes* 23: 709-748.

Kaiser, E., & Vihman, V. (2006). On the referential properties of Estonian pronouns and demonstratives. Paper presented at the *22nd Scandinavian Conference of Linguistics*, Aalborg, University, Denmark.

Karmiloff-Smith, A. (1986). From metaprocess to conscious access: evidence from children's metalinguistic and repair data. *Cognition* 23: 95-147.

Kraš, T., Sturt, P., & Sorace, A. (2014). Native and non-native processing of Italian subject pronouns: evidence from eye-tracking. Paper presented at *AMLaP XX*, Edinburgh 4-6 September 2014.

Laka, I. (1996). *A brief grammar of euskara, the Basque language*, Ms., University of the Basque Country, Vitoria-Gasteiz. Available online: http://www.ehu.es/grammar.

Mayol, L. (2009). *Pronouns in Catalan: information, discourse and strategy*. Unpublished doctoral dissertation. University of Pennsylvania, USA.

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

O'Grady, W. (2005). *Syntactic carpentry*: *An emergentist approach to syntax*. Mahwah, NJ: Lawrence Erlbaum.

Papadopoulou, D., Peristeri, E., Plemenou, E., Marinis, T., & Tsimpli, I. (2015). Pronoun ambiguity resolution in Greek: Evidence from monolingual adults and children. *Lingua*: 98–120.

Rijk, R. P. G. de (2008). *Standard Basque. A Progressive grammar*. Cambridge, MA: MIT Press.

Serratrice, L., Sorace, A., & Paoli, S. (2004). Crosslinguistic influence at the syntax pragmatics interface: Subjects and objects in English-Italian bilingual and monolingual acquisition. *Bilingualism: Language and Cognition* 7: 183-205.

Shin, N.L. (2015). Acquiring constraints on morphosyntactic variation: Children's Spanish subject pronoun expression. *Journal of Child Language*. http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=9883308&fileId=S0305000915000380

Shin, N. L. & Cairns, H. S. (2012). The development of NP selection in school-age children: reference and Spanish subject pronouns. *Language Acquisition* 19: 3-38.

Sorace, A. (2000). Differential effects of attrition in the L1 syntax of L2 near-native speakers. In In S. C. Howell, S. A. Fish, and T. Keith-Lucas (eds.), *Proceedings of the 24th Boston University Conference on Language Development (BUCLD24)*, pp. 719-725. Somerville, MA: Cascadilla Press.

Sorace, A. (2011). Pinning down the concept of 'interface' in bilingualism. *Linguistic Approaches to Bilingualism* 1: 1-33.

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research* 22: 339-368.

Sorace, A., & Serratrice, L. (2009). Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism* 13: 195-210.

Sorace, A., Serratrice, L., Filiaci, F., & Baldo, M. (2009). Discourse conditions of subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua* 119: 460-477.

Wilson, F. (2009). *Processing at the syntax-discourse interface in second language acquisition.* Unpublished doctoral dissertation. University of Edinburgh, UK.