

Bernhard Brehmer*, Aldona Sopata and Massimiliano Canzi

Same or different? Subject realization in the majority and the heritage language of Polish-German bilingual children

<https://doi.org/10.1515/lingvan-2022-0061>

Received May 16, 2022; accepted November 10, 2022; published online March 22, 2023

Abstract: The paper examines the extent to which bilingual children select lexical noun phrases and null and overt pronouns as referring expressions in their majority language German and their heritage language Polish. Both languages are similar regarding the availability of lexical noun phrases but differ in terms of the distribution of null and overt pronominal forms. Our focus lies on discourse contexts with a subject antecedent in the preceding clause, which require only light processing for both speaker and hearer due to the high accessibility of the intended subject referent. Drawing on experimental data from a picture story retelling task (MAIN) to investigate the distribution of referring expressions in the two languages compared to age-matched monolingual control groups, our results reveal that bilingual children are sensitive to crosslinguistic differences in the syntactic and discourse-pragmatic constraints that regulate the distribution of null and overt subjects in Polish and German, depending on the mode of speech (narrative or dialogic). Furthermore, there are no significant differences between the bilingual and monolingual children, irrespective of language and age group. Thus, our study cannot confirm findings of previous studies concerning the tendency of bilingual children to be either overspecific or underspecific in subject reference production.

Keywords: bilingual reference production; heritage language; lexical NPs; majority language; null subjects; pronouns

1 Introduction

The knowledge about different types of referential expressions (henceforth REs) like full noun phrases (NPs) or pronouns that can be used to refer to entities in the surrounding world represents a crucial aspect of language comprehension and production. Even for monolingual children the selection of the most appropriate expression to refer to a given discourse entity poses a huge challenge. The situation gets even more complicated for bilingual children: they have to acquire reference production and understanding in two languages where certain REs can occur in both languages (albeit, possibly, with a different distribution), while other options are available in only one of them. This might result in crosslinguistic influence, when language-specific devices or rules are over-applied and transferred to the bilingual's other language.

In the current study we analyze subject REs in a picture story task performed in both the majority language (ML) German and the heritage language (HL) Polish by bilingual children growing up in Germany. Polish represents a canonical null-subject language, while German normally tends to use overt subjects (see Section 3 for more details). Our main aim is to provide evidence of which factors contribute to possible variations in subject reference use by highlighting the effect of chronological age and mode of speech (narrative vs. dialogic) on referential choice in bilinguals. Furthermore, we look at the relevance of crosslinguistic influence (CLI) in argument realization, which has been controversially debated in previous research.

*Corresponding author: **Bernhard Brehmer**, University of Konstanz, Konstanz, Germany, E-mail: bernhard.brehmer@uni-konstanz.de
Aldona Sopata, Adam Mickiewicz University Poznań, Poznań, Poland, E-mail: sopata@amu.edu.pl
Massimiliano Canzi, University of Konstanz, Konstanz, Germany, E-mail: massimiliano.canzi@uni-konstanz.de

2 Subject argument realization in bilingual children

The comprehension and production of REs in bilinguals has been the topic of a very large number of studies over the previous decades (e.g., Contemori and Ivanova 2021; Serratrice and Hervé 2015; Torregrossa et al. 2021). However, studies differ greatly with regard to the type of bilingual speaker (L2 learners, [early] heritage bilinguals, L1 attriters), the investigated age groups (children vs. adults), the type of data used, and the language pairs, as well as the factors that are looked at as potentially influencing the distribution of REs. The latter can be roughly divided into purely linguistic factors, discourse-related factors, language-external factors, and cognitive factors. Therefore, the results of these studies are not always comparable, and this might account for the inconclusive evidence provided by them. Many studies have found evidence for an overproduction of explicit subject expressions in null-subject languages in contact with non-null-subject languages, especially in topic-continuity contexts, and treated this outcome as indicative for CLI (e.g., Otwinowska et al. 2022). However, other studies have found the same effect when investigating individuals acquiring two null-subject languages (e.g., Georgopoulos 2017; Lozano 2006). These findings led to the assumption that the tendency to be overspecific by introducing overt subjects where a null element would suffice is a general phenomenon of bilinguals, irrespective of the subject property status of the languages in contact. According to the very often cited Interface Hypothesis (see Sorace 2011 for a comprehensive overview), the choice of REs poses a challenge to bilinguals because it “involves both mastering of language-specific syntax-to-discourse mappings as well as cognitive processes such as (discourse) updating” (Torregrossa et al. 2021: 694). Even in non-null-subject languages, bilinguals can differ from monolinguals in their preferences for certain REs (Contemori and Dussias 2016).

Previous research has shown that the chronological age of bilinguals seems to play an important role in target-like mastery of the use of REs, and this also relates to the role of cognitive capacity in bilinguals' referential choice. In a series of studies on the use of null subjects in Romance HLs in contact with English or German, age turned out to be a relevant factor: While children between the ages of 1;6 and 5;0 who grow up simultaneously with Italian as a HL and English or German as MLs showed a clear tendency to overproduce overt subjects in heritage Italian (Schmitz 2007; Schmitz et al. 2012; Serratrice 2007a; Serratrice and Sorace 2002; Serratrice et al. 2004), older simultaneous bilingual children between the ages of 5;0 and 8;0 did not (Serratrice 2007a; Wolleb 2013). The same applied to adult heritage speakers (Kupisch 2012; Nagy 2015; Nagy et al. 2011; Schmitz et al. 2016). These findings highlight the importance of language experience for the acquisition of REs in bilinguals (Torregrossa et al. 2021).

In our study, we address the role of chronological age on the selection of subject reference choice by drawing on data from an under-studied population, namely Polish-German bilingual children. Data from Slavic languages other than Russian (which is, unlike Polish, not a prototypical null-subject language; see Ivanova-Sullivan 2014) have seldom been used to test the effect of the factors described above on the selection of subject expressions in bilingual contexts (see Otwinowska et al. 2022 for one study on Polish-English bilingual children).

3 Subject realization in Polish and German

The general inventory of subject REs includes full NPs, overt pronouns, and also null pronouns in both languages. Full NPs are required in Polish and German for referents with a low degree of syntactic or discourse accessibility, that is, when more explicit information is needed in order to identify the referent. Therefore, NPs normally occur in contexts that serve to introduce new referents and topics or when a referent is taken up that has not been mentioned in the previous discourse. Overt pronouns (mostly personal pronouns but also demonstratives) are used in both languages to refer to antecedents with a higher degree of activation¹ in the

¹ Following Serratrice (2007b: 185), we define an active referent as “one that is salient in the hearer's consciousness at a particular moment in time. An active referent is highly accessible either through recent prior mention, or through being the focus of the speaker's and hearer's joint attention.”

preceding discourse in cases where there is a need for conveying emphasis or to avoid potential ambiguities. German and Polish differ, however, with regard to the distribution of phonetically non-realized (= null) subjects. In Polish, null pronouns are the default option in the case of a high accessibility of the referent. Moreover, Polish null subjects are genuine null arguments of the pro-drop type conditioned by subject-verb agreement (see Sigurðsson 2011). Thus, null pronouns as the maximally reduced form of REs are the preferred option for salient subject antecedents in Polish, while less salient non-subject antecedents favor forms that will more explicitly signal the shift of the topic, for example, overt pronouns. Therefore, we find a complementary distribution of null and overt subject pronouns dependent on the status of the antecedent in Polish: null subjects are the pragmatically optimal option when they are co-referential with a topic antecedent. In German, independent of both syntactic and discourse-pragmatic constraints, the default option will be an explicit subject expression (pronoun or NP, depending on the saliency of the referent). Nevertheless, subjects can be dropped in German, too. These instances are restricted to the topic position of the empty element (“topic drop”; see Grewendorf 1989). In both languages, empty elements in a subject position appear as a rule in root contexts, for example in coordinate clauses, as in (1).²

- (1) a. German: *Er kam nach Hause und Ø öffnete die Tür.*
 he come.PST to house and open.PST the door
 b. Polish: *On wrócił do domu i Ø otworzył drzwi.*
 he come.PST to house.GEN and open.PST door
 ‘He came home and Ø opened the door.’

4 The study

4.1 Research questions

Our main goal is to assess whether Polish-German bilingual children behave differently in their two languages when it comes to the type of REs used to denote subject arguments and whether their choice resembles the distribution found in monolingual data for both languages. Thus, we address the following research questions (RQs):

- RQ 1: Is the distribution of REs coding subject arguments the same for German and Polish in the Polish-German bilingual group?
- RQ 2: Does the distribution of subject REs in the bilingual data differ from preferences found in two age-matched monolingual control groups?
- RQ 3: Is the distribution of subject REs dependent on the chronological age of the bilingual children and the mode of speech (narrative vs. dialogic)?

By collecting and analyzing evidence from an under-studied language pair, our study contributes to the state of the art of the field on the basis of novel empirical data. We can compare our results to previous studies dealing with issues of subject reference choice in bilinguals and shed new light on both the relevance and directionality of CLI and the factors contributing to possible deviations from preferences for specific types of REs found in monolinguals. As has been shown in Section 1, previous research has established a whole range of factors that proved to be influential for subject reference choices. To produce meaningful results, we decided to control for as many factors as possible and focus mostly on the role of (a) the bilinguals’ chronological age and (b) the speech mode (narrative vs. dialogic).

² This example follows the Leipzig Glossing Rules. Abbreviations used: GEN genitive; PST past tense.

4.2 Participants

To control for language experience and quantity of input, we focused on children who grew up simultaneously with both the ML German and the HL Polish.³ Following a proposal of Meisel (2007), we consider bilingual children who started to acquire both languages before the age of 3 as simultaneous bilinguals. Overall, our data come from 32 bilingual Polish-German children living in Germany. To establish our control groups, we collected data from 26 age-matched monolingual German children living in Germany and from 20 monolingual Polish children living in Poland. Both monolinguals and bilinguals were divided into two age groups depending on their chronological age at testing. The cut-off point between the two age groups was set at the age of 8, as this resulted in two groups of roughly equal size. Consequently, our sample consists of six groups (see Table 1).

4.3 Data collection

Data were gathered through an elicited oral production task. We used the Multilingual Assessment Instrument for Narratives (MAIN) designed by Gagarina et al. (2012) and chose two of the four picture stories (“Baby Birds” and “Baby Goats”). First, the children had to produce a coherent narrative based on the picture stories. After that, the experimenter asked a fixed set of questions about the protagonists of the stories and their possible motivations for acting the way they did in the story. The latter part constituted the dialogic subcorpus of our study and contained the children’s responses to the questions posed by the adult experimenter. The narrative and the dialogic parts were analyzed separately.

4.4 Data analysis

To control for as many factors as possible, we did not analyze all relevant clauses in our sample but focused on a subset. Apart from language experience (see Section 4.2) we controlled for the following factors:

Table 1: Participant groups in the study, together with their sex, age, and age of onset (AoO) of Polish and German acquisition.

Group	N	Sex	Age (mean/span)	AoO Polish	AoO German (mean/span)
Younger bilinguals	16	10 males 6 females	6;5/4;11–8;0	0	0;11 0–2;11
Older bilinguals	16	9 males 7 females	11;8/8;3–13;11	0	1;3 0–2;11
Younger monolinguals (German)	11	5 males 6 females	5;11/4;4–7;6	–	0
Older monolinguals (German)	15	10 males 5 females	9;10/8;1–13;6	–	0
Younger monolinguals (Polish)	9	4 males 5 females	5;7/3;7–8;0	0	–
Older monolinguals (Polish)	11	7 males 4 females	9;7/8;3–11;4	0	–

³ The term “heritage speaker” denotes “an early bilingual who grew up hearing (and speaking) the heritage language (L1) and the majority language (L2) either simultaneously or sequentially in early childhood (that is, roughly up to age 5), but for whom L2 became the primary language at some point during childhood (at, around, or after the onset of schooling)” (Benmamoun et al. 2013: 133). It refers to a specific sociolinguistic setting of bilingual language acquisition and does not depend on the level of proficiency in both languages; that is, it does not exclude balanced bilinguals, who can be found in heritage speaker populations as well (see Kupisch 2013). The situation of Polish as a minority language in Germany justifies its application to the current sample, although the geographical circumstances allow for frequent visits to the country of origin.

- (a) *Referent accessibility*: We included only those clauses in the analysis where the subject referent is identical to the subject referent of the preceding clause. Thus, we analyzed only instances of topic continuity and referent maintenance where the antecedent occupies a prominent syntactic position (subject of the preceding clause) and is therefore highly activated and salient from a discourse-pragmatic perspective as well.
- (b) *Grammatical person*: While first and second person referents are easily accessible as they identify the speaker or the addressee of the utterance, third person forms are potentially ambiguous as they can refer to a wider range of potential referents. For this reason, we included only third person referents in our analysis.
- (c) *Animacy*: As all protagonists of the MAIN picture stories are animals (birds, goats, a fox, a dog, a cat), we could control for the effect of animacy of the subjects (Torregrossa et al. 2021: 699).

As stated above, we analyzed the tokens gained from the narrative and dialogic parts of the test separately. Table 2 shows the number of tokens included in the final analysis and their percentage of the total number of clauses in the analyzed corpus.⁴

The data show that subject referent maintenance occurs frequently in the narrative as well as in the dialogic parts of the analyzed texts. Its relative frequency is usually higher in the dialogues than in the narratives, and this holds for both languages.

5 Results

Subject reference choices were first analyzed descriptively by presenting the relative frequency of the following types of REs in the children's data: lexical NPs, personal pronouns (PP), demonstratives (DP), and null pronouns (null). We compared the distribution of these types of REs in the German and Polish data of the two bilingual age groups (see Figure 1). Data from the narrative (left graph) and the dialogic parts (right graph) are given separately.

Figure 1 shows that lexical NPs are more frequent in the narrative parts, where they appear in both languages and both age groups to a comparable degree (15–20%). In the dialogic parts, this maximally informative type of RE is the least frequently chosen option by both age groups in both of their languages. Explicit personal pronouns are surprisingly infrequent in the German narratives of both age groups, while they remain marginal in the Polish data, which was expected due to the null-subject property of Polish. However, in the dialogic part they are clearly the preferred option in German. Here, pronouns (personal and demonstrative) comprise between 92% (younger group) and 75% (older group) of all REs used. In the Polish data, null pronouns clearly prevail irrespective of text type and age group (between 75 and 89%), but even in the German narratives they make up at least half of all

Table 2: Number of tokens included in the analysis and the percentage of total clauses that contain subject reference maintenance.

Group	Language	Tokens in narrative part (and as percentage of total clauses)	Tokens in dialogic part (and as percentage of total clauses)
Younger bilinguals	German	70 (32)	79 (53)
Older bilinguals	German	86 (32)	149 (60)
Younger bilinguals	Polish	55 (22)	99 (57)
Older bilinguals	Polish	84 (33)	189 (67)
Younger monolinguals (German)	German	55 (40)	40 (39)
Older monolinguals (German)	German	83 (35)	122 (60)
Younger monolinguals (Polish)	Polish	37 (30)	61 (56)
Older monolinguals (Polish)	Polish	47 (31)	113 (67)

⁴ We excluded from the analysis those clauses with impersonal constructions and imperative phrases, and clauses with relative or indefinite pronouns or numerals in subject position.

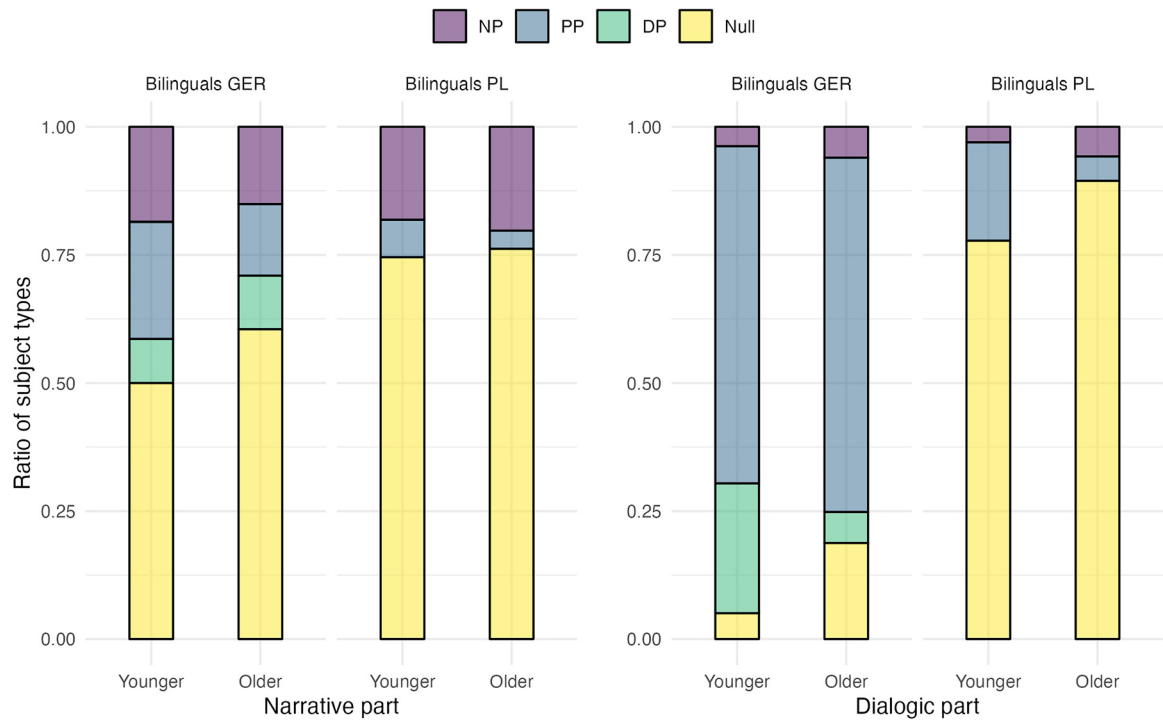


Figure 1: Relative frequency of subject reference types (full NPs, personal pronouns, demonstratives, and null pronouns) in the German and the Polish data of bilingual children in the narrative (*left*) and dialogic parts (*right*).

recorded tokens. This stands in stark contrast to the dialogic parts where null pronouns in German are marginal, especially in the younger group of bilinguals.

Figure 2 introduces the distribution of subject REs in German by contrasting the data gathered from the bilinguals and an age-matched monolingual control group. The distribution of REs again shows considerable differences depending on the text type. For both bilingual and monolingual age groups, null subjects are the most frequent type of subject expression in narrative contexts, which is rather surprising given that German is a topic drop language, where explicit pronouns normally represent the default option in case of a high accessibility of the subject referent. Consequently, demonstrative and personal pronouns occupy the second position when it comes to frequency, with a higher percentage in the younger bilingual and monolingual groups. The distribution of the types of REs is very similar among the bilingual and monolingual groups. The data from the dialogic part exhibit a distribution that is clearly different from the narrative text passages. Here, both types of pronouns function together as the default means of subject referent maintenance. Demonstratives are more frequent in the two younger age groups and lose importance with increasing age. Lexical NPs are hardly used when children can take up the subject referent that was mentioned by the adult interlocutor in a preceding question. In the narrative parts, they play a more important role. Although the distribution of REs in the dialogic part deviates from the one found in the narrative mode, there are no clear differences between the corresponding bilingual and monolingual age groups.

The picture changes when we compare subject realization in Polish narratives of the bilingual groups and age-matched monolingual groups (see Figure 3). Both bilingual age groups use more NPs and less null subjects than their monolingual Polish peers. Overt personal pronouns occur only in the two bilingual groups. This differs from the data in the dialogic parts, where personal pronouns are used by both bilingual and monolingual age groups. However, null subjects outnumber the instances where overt subject REs are used in this text type by far, which parallels the findings for the narrative parts. In contrast to the narrative mode, however, the distribution in the dialogic data does not exhibit huge differences between the bilingual and monolingual groups.

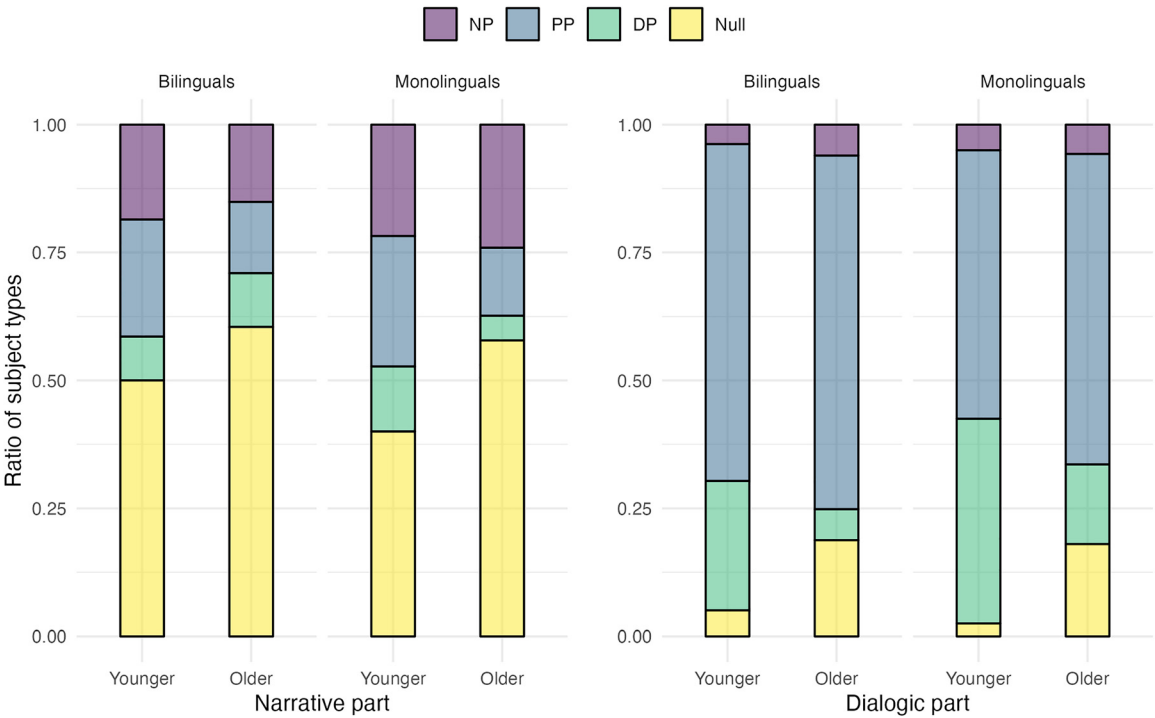


Figure 2: Relative frequency of subject reference types (full NPs, personal pronouns, demonstratives, and null pronouns) in the German data of bilingual and monolingual children in the narrative (*left*) and dialogic parts (*right*).

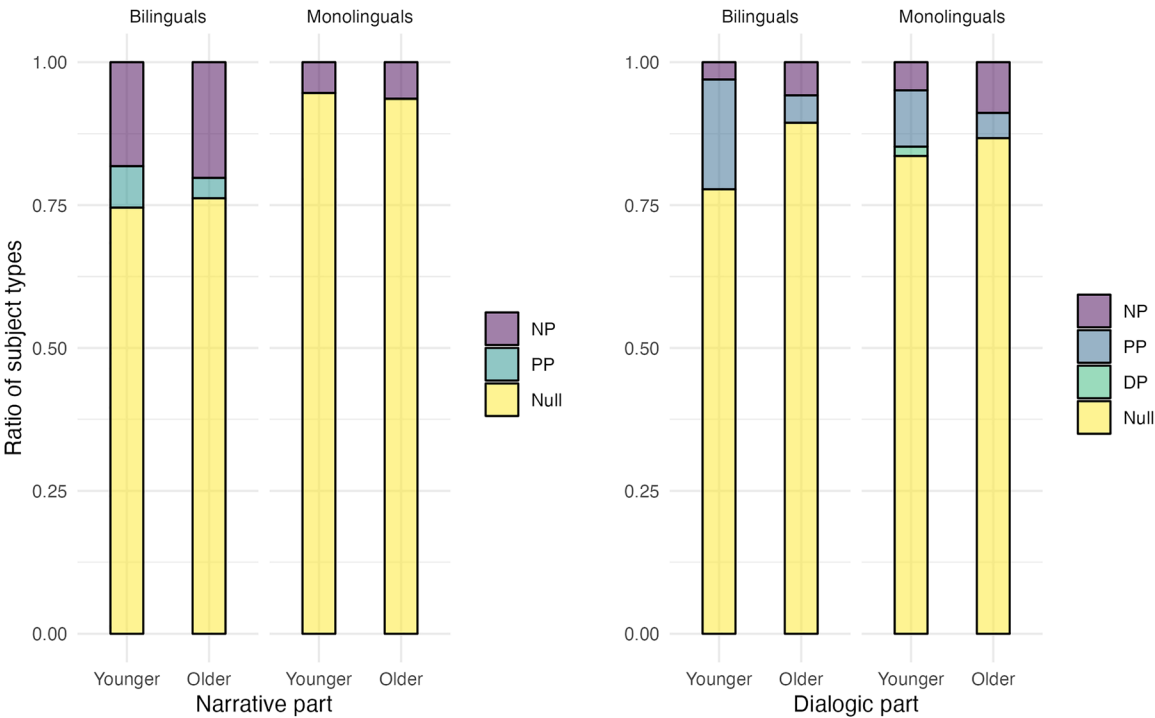


Figure 3: Relative frequency of subject reference types (full NPs, personal pronouns, demonstratives, and null pronouns) in the Polish data of bilingual and monolingual children in the narrative (*left*) and dialogic parts (*right*).

To check whether the differences that were found in the descriptive analysis reach statistical significance, we conducted an analysis employing the R statistical language (version 4.0.3; R Core Team 2020) on macOS Big Sur 10.16, where we used the packages lme4 (version 1.1.26; Bates et al. 2015), lmerTest (version 3.1.3; Kuznetsova et al. 2017), emmeans (version 1.5.4; Lenth 2021), and car (version 3.0.10; Fox and Weisberg 2019). For each language, we fit a mixed Poisson regression model (estimated using maximum likelihood and BOBYQA optimizer) to predict counts of subject with the interaction of subject type (“NP”, “DP”, “PP”, and “Null”) and age group (“Older”, “Younger”), subject type and bilingualism status (“Bilingual”, “Monolingual”), and subject type and text type (“Dialogue”, “Narrative”). The models included varying intercepts for participants, to account for differences in overall subject counts. A slope for passage type over participant was originally added to both models but was later removed when the models resulted in a singular fit. For Polish, none of the interactions of interest were significant at an alpha threshold of 0.05, suggesting that differences in age, bilingualism status, and text type did not significantly predict distributional changes of subject type counts across participants. For German, only the interaction between subject type and text type was reported as significant ($\chi^2(3, N = 254) = 46.42, p < 0.001$). Post hoc pairwise contrasts carried out in emmeans suggest that participants use significantly more DP ($z = 2.20, p < 0.05$) and PP ($z = 6.55, p < 0.0001$) in the dialogues compared to the narratives, while null subjects are significantly more predominant ($z = -2.67, p < 0.01$) in the narrative task. However, interactions of subject type and bilingualism status and age group were not significant, suggesting that neither significantly predicts distributional changes in subject type. Finally, we fit a mixed Poisson regression (estimated using maximum likelihood and BOBYQA optimizer) to predict estimates of subject counts of bilingual participants only, to compare their performance in both German and Polish. The fixed structure of the model included a three-way interaction of subject type, L1 (“German”, “Polish”), and text type. Varying intercepts were allowed for participants. A slope for passage type was originally included but later removed when the regression model resulted in an over-fitted singular fit. The three-way interaction of L1, text type, and subject type was significant ($\chi^2(2, N = 164) = 6.0434, p < 0.0005$). Post hoc pairwise contrasts suggest that, for dialogues, bilingual participants are significantly more likely to use “PP” ($z = 3.865, p < 0.0005$) in German compared to Polish. In contrast, participants are significantly less likely ($z = -7.584, p < 0.0005$) to use “Null” in German, compared to Polish, given the same type of text.

6 Discussion

6.1 Subject realization in the majority and the heritage language of bilinguals (RQ 1)

The statistical analysis revealed a significant difference in subject reference choice between the bilingual children’s ML German and HL Polish. However, this difference depends on the type of speech mode (narrative vs. dialogic). While for narrative contexts we could not prove a difference in subject REs between Polish and German, there is a significant difference when it comes to the dialogic parts. Here, both the younger and the older bilingual children use overt (personal and demonstrative) pronouns to a significantly higher extent in German (92 and 75%) than in Polish, where in turn null subjects are clearly the preferred option (78 and 89%; see Figure 1). This is in line with the predictions of models on the influence of referent accessibility on the choice of subject REs and the divergent “default” forms in German and Polish: while in German overt subject pronouns are used not only to indicate topic shift, contrast, or emphasis but also serve to express the subject in neutral contexts or even to maintain topic continuity, null subjects perform the latter function in Polish (see Section 3). Thus, bilingual children are sensitive to crosslinguistic differences in the syntactic and discourse-pragmatic constraints that regulate the distribution of null and overt subjects. The question arises, however, why the bilinguals show a similar distribution of REs in German and Polish in the narrative parts of the picture storytelling. The high number of phonetically non-realized subjects in German is due to the fact that in the narrative parts the children produce longer utterances including coordinate clauses, especially in German, the language in which the children are exposed to more input from outside the family, as it is the ML that dominates their surroundings. It is in these coordinated clauses where subject ellipsis mainly occurs, because it represents the only pragmatically felicitous option in German (see example (1a)). However, one must not forget that German is a language which allows for

topic drop, especially in colloquial language use. This accounts for the occurrence of null subjects in the dialogic parts of the German texts as well, especially in the older age group (see Figure 1). The preference for overt pronouns in the dialogic parts of the German texts can be explained by the fact that the children react to the immediately preceding questions of the experimenter about the behavior and the motivations behind the actions of the animal characters in the picture stories. This favors the use of explicit pronouns at the beginning of their answers. In Polish, due to subject-verb agreement rules, null subjects are the preferred option in these cases, and this is mirrored by the behavior of our bilingual children in their HL.

6.2 Subject realization in bilingual and monolingual children (RQ 2)

Statistical testing for the effect of bilingualism on the distribution of subject argument realization revealed no significant differences between the bilingual and monolingual children, irrespective of the language. This holds for both the narrative and the dialogic parts.

Consequently, our study cannot confirm findings of previous studies where differences between monolinguals and bilinguals were found and where these differences were accounted for either by CLI or by general processing problems of the bilinguals, as predicted by the Interface Hypothesis (see Section 2). While the similar distribution of subject REs in bilinguals and monolinguals in the ML German might be expected, the failure to find an effect of bilingual language experience on variation in the use of subject REs in the (presumably weaker) HL Polish calls for an explanation. Previous research reported on an overuse of overt pronominal subjects in contexts in which a null subject would be pragmatically appropriate (which also applies to Polish as a canonical null-subject language; see Contemori and Ivanova 2021 for an overview). This effect emerges predominantly in contexts with subject referent maintenance or topic continuity, as investigated in the current study. Instances where a speaker refers to a highly activated referent by using an explicit RE even though a reduced form would suffice are treated as examples of overspecification in reference production (Torregrossa et al. 2021). Our findings show that the acquisition of a null-subject HL does not necessarily lead to overspecification in reference production. The result that the Polish-German bilinguals in our study did not overuse overt subject pronouns in Polish may be due to the fact that the children's other language, German, normally uses overt subjects but does also allow for phonetically empty subjects in topic drop constructions. This might explain the contrast between the bilinguals in the present study and the Polish-English bilinguals investigated by Otwinowska et al. (2022), who overused overt pronouns in Polish. Since English is a canonical non-null-subject language, transfer of the overt subject property resulting in overspecification in reference production in Polish is more likely to happen there than in the case of Polish-German bilinguals.

While we did not find an overextension of overt subject pronouns, we did encounter a slight tendency in the bilinguals to use more lexical NPs than their monolingual peers, at least in the narrative data (18–20% vs. 5–6%; see Figure 3; see also Montrul 2004; Ryan 2015; Torregrossa and Bongartz 2018 for similar findings). We treat this trend (which did not reach significance) as a reflex of a stronger orientation towards audience design in bilinguals when compared to monolinguals. Since NPs function as the maximally informative way of identifying subject referents compared to pronouns, linguistic insecurity in the weaker HL might result in preferring this type of RE to get the meaning across to the interlocutor.

6.3 Age and speech mode effects (RQ 3)

Chronological age did not emerge as a predictor for changes in the distribution of subject REs in the bilingual and monolingual data. At least for children older than 5 years, there are no significant differences between bilinguals and monolinguals regarding preferred options for realizing subject arguments. This corroborates findings from previous studies on children with Romance HLs (e.g., Serratrice 2007a; Wolleb 2013).

Speech mode, however, served as a significant predictor for the selection of REs, at least in the German data. In the dialogic parts, bilingual and monolingual German children show the reverse preference for overt subject

pronouns compared to the narrative parts where null subjects prevail. This difference is probably due to the higher activation of the referent on the part of the hearer in the case of the dialogic question-answer section. It reduces the cognitive load for the children. Thus, they can more reliably count on the recoverability of the subject referent which was previously mentioned by the interlocutor him/herself, which leads to a preference of the default form (i.e., PP in German and null subject in Polish).

7 Conclusions

Our study on subject REs in Polish-German bilingual children could not confirm findings from previous studies that bilingual language experience is a relevant factor that contributes to variation in the use of REs. Although the language pair of Polish (a canonical null-subject language) and German (a topic drop language that nevertheless favors overt pronouns or lexical NPs to mark the subject of a sentence) shares the basic structural characteristics of language combinations that have been investigated in previous studies, we found neither a general tendency towards overspecification (e.g., the use of an overt personal pronoun where a null subject would suffice) nor an overuse of underspecified REs (e.g., the choice of ambiguous pronouns instead of more explicit lexical NPs) in our bilingual data compared to the monolinguals. While the bilinguals exhibited a stronger trend towards the use of lexical NPs in instances where a personal or null pronoun would be the more economical option in the narrative, this quantitative difference turned out not to be significant. Our study of Polish-German bilingual children between the ages of 5 and 13 shows that simultaneous bilinguals do not exhibit categorical differences in the distribution of subject REs when compared to age-matched monolinguals in both of their languages. The fact that we were not able to find effects of chronological age, potential processing difficulties in bilinguals (as predicted by the Interface Hypothesis), or CLI may be due to the restriction of only focusing on instances of subject referent maintenance (i.e., topic continuity); that is, we considered only cases where the referent is highly accessible both from a syntactic (subject antecedent) and discourse-pragmatic view (antecedent occurs in the preceding clause). These cases, therefore, require only light processing by the speaker (and hearer), which might have facilitated the selection of language-specific, discourse-appropriate REs in the realization of subject arguments. Our findings underline the fact that children growing up simultaneously with two languages can indeed exhibit target-like representations of the system of REs in both of their L1s and can cope with the cognitive complexity that requires the matching of syntactic and discourse information. We suppose that extralinguistic/social factors (such as a relatively high amount of language exposure and use, especially in the HL Polish) play a crucial role in the acquisition of target-like representations for both languages (see also Meisel 2011 for a general account of this issue).

Research funding: This work was supported by Narodowe Centrum Nauki (2014/15/G/HS6/04521) and Deutsche Forschungsgemeinschaft (277135691).

References

- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.
- Benmamoun, Elabbas, Silvina Montrul & Maria Polinsky. 2013. Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics* 39(3–4). 129–181.
- Contemori, Carla & Paola E. Dussias. 2016. Referential choice in a second language: Evidence for a listener-oriented approach. *Language, Cognition and Neuroscience* 31(10). 1257–1272.
- Contemori, Carla & Iva Ivanova. 2021. Bilingual referential choice in cognitively demanding situations. *Bilingualism: Language and Cognition* 24(1). 83–95.
- Fox, John & Sanford Weisberg. 2019. *An R companion to applied regression*, 3rd edn. Thousand Oaks, CA: Sage. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/> (accessed 19 April 2022).

- Gagarina, Natalia, Daleen Klop, Sari Kunnari, Koula Tantele, Taina Välimaa, Ingrida Balčiūnienė, Ute Bohnacker & Joel Walters. 2012. *MAIN: Multilingual assessment instrument for narratives*. Berlin: ZAS. <https://d-nb.info/1061072533/34> (accessed 19 April 2022).
- Georgopoulos, Athanasios. 2017. *Anaphora in the interlanguage of English and Greek learners of L2 Spanish: A study based on the CEDEL2 corpus*. Granada: Universidad de Granada doctoral dissertation.
- Grewendorf, Günther. 1989. Small pro in German. In Günther Grewendorf & Wolfgang Sternefeld (eds.), *Scrambling and barriers*, 294–315. Amsterdam: John Benjamins.
- Ivanova-Sullivan, Tanya. 2014. Interpretation of anaphoric subject pronouns in heritage Russian. In Ryan T. Miller, Katherine I. Martin, Chelsea M. Eddington, Ashlie Henery, Nausica Marcos Miguel, Alison M. Tseng, Alba Tuninetti & Daniel Walter (eds.), *Selected proceedings of the 2012 second language research forum*, 145–154. Somerville, MA: Cascadia Proceedings Project.
- Kupisch, Tanja. 2012. Specific and generic subjects in the Italian of German-Italian simultaneous bilinguals and L2 learners. *Bilingualism: Language and Cognition* 15(4). 736–756.
- Kupisch, Tanja. 2013. A new term for a better distinction? A view from the higher end of the proficiency scale. *Theoretical Linguistics* 39(3–4). 203–214.
- Kuznetsova, Alexandra, Per B. Brockhoff & Rune Haubo Bojesen Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13). 1–26.
- Lenth, Russell V. 2021. emmeans: Estimated marginal means, aka least-squares means, version 1.5.4 [R package]. <https://CRAN.R-project.org/package=emmeans> (accessed 19 April 2022).
- Lozano, Cristóbal. 2006. The development of the syntax-discourse interface: Greek learners of Spanish. In Vincent Torrens & Linda Escobar (eds.), *The acquisition of syntax in Romance languages*, 371–399. Amsterdam: John Benjamins.
- Meisel, Jürgen M. 2007. The weaker language in early child bilingualism: Acquiring a first language as a second language? *Applied Psycholinguistics* 28(3). 495–514.
- Meisel, Jürgen M. 2011. *First and second language acquisition: Parallels and differences*. Cambridge: Cambridge University Press.
- Montrul, Silvina. 2004. Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition* 7(2). 125–142.
- Nagy, Naomi. 2015. A sociolinguistic view of null subjects and VOT in Toronto heritage languages. *Lingua* 164B. 309–327.
- Nagy, Naomi, Nina Aghdasi, Derek Denis & Alexandra Motut. 2011. Null subjects in heritage languages: Contact effects in a cross-linguistic context. *University of Pennsylvania Working Papers in Linguistics* 17(2). 16.
- Otwinowska, Agnieszka, Marcin Opacki, Karolina Mieszkowska, Marta Białecka-Pikul, Zofia Wodniecka & Ewa Haman. 2022. Polish-English bilingual children overuse referential markers: MLU inflation in Polish-language narratives. *First Language* 42(2). 191–215.
- R Core Team. 2020. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/> (accessed 19 April 2022).
- Ryan, Jonathon. 2015. Overexplicit referent tracking in L2 English: Strategy, avoidance, or myth? *Language Learning* 65(4). 824–859.
- Schmitz, Katrin. 2007. L'interface syntaxe-pragmatique: Le sujet chez des enfants franco-allemands et italo-allemands [The syntax-pragmatics interface: The subject in French-German and Italian-German children]. *Acquisition et Interaction en Langue Etrangère* 25. 9–43.
- Schmitz, Katrin, Laura Di Venanzio & Anna-Lena Scherger. 2016. Null and overt subjects in Italian and Spanish heritage speakers in Germany. *Lingua* 180. 101–123.
- Schmitz, Katrin, Marisa Patuto & Natascha Müller. 2012. The null-subject parameter at the interface between syntax and pragmatics: Evidence from bilingual German-Italian, German-French and Italian-French children. *First Language* 32(1–2). 205–238.
- Serratrice, Ludovica. 2007a. Cross-linguistic influence in the interpretation of anaphoric and cataphoric pronouns in English-Italian bilingual children. *Bilingualism: Language and Cognition* 10(3). 225–238.
- Serratrice, Ludovica. 2007b. Null and overt subjects at the discourse interface: Evidence from monolingual and bilingual acquisition. In Sergio Baauw, Jacqueline van Kampen & Manuela Pinto (eds.), *The acquisition of Romance languages*, 181–200. Utrecht: LOT.
- Serratrice, Ludovica & Antonella Sorace. 2002. Overt and null subjects in monolingual and bilingual Italian acquisition. In Barbara Beachley, Amanda Brown & Frances Conlin (eds.), *BUCLD 27: Proceedings of the 27th annual Boston University conference on child language development*, 739–750. Somerville, MA: Cascadia Press.
- Serratrice, Ludovica, Antonella Sorace & Sandra Paoli. 2004. Cross-linguistic influence at the syntax-pragmatics interface: Subjects and objects in English-Italian bilingual and monolingual acquisition. *Bilingualism: Language and Cognition* 7(3). 183–205.
- Serratrice, Ludovica & Coralie Hervé. 2015. Referential expressions in bilingual acquisition. In Ludovica Serratrice & Shanley Allen (eds.), *The acquisition of reference*, 311–333. Amsterdam: John Benjamins.
- Sigurðsson, Halldór Ármann. 2011. Conditions on argument drop. *Linguistic Inquiry* 42(2). 267–304.
- Sorace, Antonella. 2011. Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism* 1(1). 1–33.
- Torregrossa, Jacopo & Christiane Bongartz. 2018. Teasing apart the effects of dominance, transfer, and processing in reference production by German-Italian bilingual adolescents. *Languages* 3(3). 36.
- Torregrossa, Jacopo, Maria Andreou, Christiane Bongartz & Ianthi Maria Tsimpli. 2021. Bilingual acquisition of reference: The role of language experience, executive functions and cross-linguistic effects. *Bilingualism: Language and Cognition* 24(4). 694–706.
- Wolleb, Anna. 2013. Referring expressions in the narratives of Italian-English bilingual children. *Studia Linguistica* 67(1). 28–46.