Referential expressions in bilingual acquisition

Ludovica Serratrice* and Coralie Hervé*,#

* The University of Manchester / # Université Paris Diderot – Paris 7

Just like monolingual children, bilingual children need to carve up the referential space to understand and produce discourse-appropriate referential expressions. In the case of bilinguals, this demanding task additionally requires language-specific form-function mappings that may be structurally similar or different in their two languages. Cases of partial form-function overlap across languages, especially with respect to third person pronouns, have been the focus of much scrutiny in connection with the issue of cross-linguistic influence. In this chapter we review naturalistic and experimental evidence showing how the degree of structural overlap across two languages, the degree of variability in the target language(s), and the amount of input that is necessary to home in on the target grammar(s) constrain the comprehension and production of referential expressions in bilingual acquisition.

Keywords: personal pronouns, cross-linguistic influence, subjects, objects, clitics

1. Introduction

Referential expressions, personal pronouns in particular, have been the focus of a substantial body of research in childhood bilingual acquisition over the last fifteen years or so.¹ Much of the interest was initially generated by the formulation of Hulk and Müller’s (2000) cross-linguistic influence (CLI) hypothesis and the subsequent emphasis on the locus of vulnerability at the syntax-pragmatics interface. The comprehension and production of referential expressions sit squarely

¹ This review deals with cases of what De Houwer (1990) defines as Bilingual First Language Acquisition (BFLA), adopting a term originally introduced by Swain (1972) and subsequently taken up by Meisel (1989). BFLA refers to cases in which children are simultaneously exposed to two languages from birth and are therefore acquiring both languages as a first language.
at the interface between syntax and pragmatics. The syntax of a language makes available a certain inventory of referential forms (e.g., the presence vs. absence of definiteness markers, null anaphors, clitics, etc.), but their distribution is constrained by the discourse-pragmatic notions of newness, joint attention and topic maintenance/shift, amongst others.

One of the key issues in the study of referential expressions in bilingual acquisition has been the realization of subject arguments, especially in pairs of languages where there is a typological asymmetry in the grammaticality of null subjects, and in the distribution of discourse-appropriate personal pronouns. Direct objects, particularly in relation to their omission rates, have also featured in the literature, although much less prominently; once again argument realization in the object domain has been mostly looked at through the prism of cross-linguistic influence.

The study of referential expressions in bilingual acquisition is still relatively recent, and the nature and the quality of the evidence are somewhat heterogeneous, making direct comparisons across different studies and disparate methodologies rather difficult. The aim of this chapter is to offer a selective review of the main studies on referential choice in bilingual children with a view to taking stock of what can now be considered as consistent and reliable patterns. In doing so, we will present the findings of research using a range of different methodologies (corpus studies, experimental tasks) and different modalities (production and comprehension), as well as focusing on different sentential arguments (subjects and objects).

2. Evidence for separate development in bilinguals’ grammar

For almost two decades in the ’80s and ’90s, the main item on the research agenda of the simultaneous acquisition of two languages was the issue of language differentiation in the domain of morphology and syntax. The systematic investigation of young bilinguals’ language use showed that children who are exposed to two languages on a regular basis from birth can acquire their two languages as independent grammatical systems, and that they can systematically use word order and morpho-syntactic markers in language-specific ways (De Houwer, 1990; Genesee, 1989; Meisel, 1989).

In the ’80s, a number of influential papers began challenging the predominant received wisdom of the time, namely, that bilingual children start out with an undifferentiated linguistic system and are unable to discriminate between phonological, lexical and morpho-syntactic aspects of the two languages (Arnberg, 1987; Leopold, 1970; Redlinger & Park, 1980; Swain, 1972; Taeschner, 1983; Toribio & Brown, 1994; Vihman, 1982, 1985; Volterra & Taeschner, 1978). On empirical
Chapter 13. Referential expressions in bilingual acquisition

and methodological grounds, Genesee (1989), Nicoladis (1994) and later Genesee, Nicoladis, and Paradis (1995) demonstrated that arguments for an undifferentiated system formulated on the basis of language mixing were ill founded. In a parallel and independent line of inquiry, evidence for the Separate Development Hypothesis (SDH) came from De Houwer’s (1990) longitudinal study of one Dutch-English bilingual child (2;7–3;4), and from findings from the DUFDE project (Meisel, 1990, 1994) including 13 bilingual children acquiring German alongside French or Italian between the ages of 18 months and 6 years.

Recently, De Houwer (2009) listed a series of studies supporting the SDH, including a total of 50 simultaneous bilingual children acquiring 17 different language combinations. Empirical findings from 21 different morpho-syntactic domains including subject realization showed that bilingual children largely do use morphology and syntax in language-appropriate ways in their two languages. Although the large evidence base for the SDH indicates that bilingual children’s two linguistic systems develop as mostly separate systems, a substantial body of research has reported the presence of pockets of cross-linguistic influence.

Hulk and Müller’s (2000) seminal CLI hypothesis suggests that this phenomenon is likely to occur in bilingual children with respect to a specific structure if (a) the structure in question is at the interface between two modules of grammar, and more particularly at the interface between discourse-pragmatics and syntax in the so-called C-domain, and if (b) there is a degree of overlap between the two language systems at the surface level (Hulk & Müller, 2000, pp. 228–229). In essence, two systems can be independent but, under certain circumstances, there may be scope for a degree of influence between them. What predicts the likelihood and the directionality of this influence has been the focus of much research over the last fifteen years (see Serratrice, 2013, for an overview). In the next two sections we will start by reviewing studies based on corpus data that speak directly to the SDH with respect to subject and object argument realization, and then turn our attention to corpus studies that have reported systematic instances of CLI.

3. Language-specific use of pronominal subjects

With specific reference to the distribution of pronominal subjects, Meisel (1989) studied two German-French bilinguals (C and P) between the ages of 1;0 and 4;0. In German, the children used subjects productively in obligatory contexts by the ages of 1;11 (C) and 2;8 (P), and until the age of 2;9 (C) and 3;3 (P) the vast majority of subjects were pronominal subjects. In French, pronominal subject clitics started appearing consistently in obligatory contexts at 1;10 (C) and 2;9 (P), and they regularly appeared in immediate pre-verbal position with or without a
left-dislocated element. Importantly, non-pronominal subjects were never found immediately adjacent to the verb in C’s speech and they did so very infrequently in P’s. Meisel (1989) argued that this reflected the distribution of subject clitics in French child directed speech, but more importantly he took it as an indication that the children respected the constraints of the target system in their two languages, in essence evidence that they treated argument forms in language-specific ways.

Paradis & Genesee (1996) also investigated the distribution of pronominal subjects in the French of three English-French bilinguals between the ages of 2;0 and 3;0. The aim of their study was to identify evidence of transfer, acceleration or delay in three specific areas where English and French differ: the distribution of pronominal subjects, the marking of finiteness, and verb placement with respect to negation. Similarly to Meisel (1989), Paradis and Genesee (1996) treated French subject pronouns more as finiteness markers than as referential expressions in their own right. Their findings also showed a consistent language-specific pattern whereby subject pronouns were only ever found with finite verb forms in French as required by the adult language (e.g., *Je peux pas dire quoi.* ‘I can’t say what.’ (Gene, 2;7)), while in English they appeared with both finite and non-finite forms (e.g., *It is mushroom.* (Gene, 2;7); *He making a fish.* (Gene, 2;7). This is yet more evidence for the SDH.

Although referential expressions featured in research on bilingual first language acquisition from the ’80s to the ’90s, it was mostly in connection with morpho-syntactic development, i.e., word order and agreement. It was not until Hulk & Müller’s (2000) and Müller & Hulk’s (2001) formulations of the CLI hypothesis that the study of referential expressions as morpho-syntactic forms mapping onto discourse functions really came into its own. This shift in focus from a mostly syntax-driven approach to a more discourse-functional approach, or at least a syntax-pragmatics interface approach, took place at the same time as the role of discourse-pragmatics in child language, especially in connection with the acquisition of reference, started gaining prominence (see Allen, Hughes, & Skarabela, this volume).

4. CLI effects in pronominal argument realization: The case of subjects

The basic task of any language learner is to map linguistic forms onto discourse functions in comprehension and production. In the case of a bilingual learner, the task is to map the same set of discourse functions onto two (potentially) distinct sets of linguistic forms. Given the underlying need for a form-function correspondence, in a bilingual context we can envisage a range of possible scenarios according to whether the repertoire of linguistic forms overlaps across the two
languages, and the extent to which the same forms map onto the same discourse functions cross-linguistically. CLI is predicted to occur in cases of partial structural overlap. In some language combinations the set of linguistic forms for a given set of discourse functions may be very similar. In the specific case of referential expressions we may have a total, or almost total, overlap in the morpho-syntactic availability of definite and indefinite articles, bare nouns, strong pronouns, clitics, and null anaphors. This is, for instance, the case of Italian and (Iberian) Spanish, two typologically related Romance languages where subject arguments can be realized by definite and indefinite NPs, and by overt and null personal pronouns; object arguments have the option of being expressed by both strong pronouns and clitics, in addition to definite and indefinite noun phrases. With specific reference to the distribution of pronominal subjects, although both languages syntactically license null and overt pronouns, according to recent evidence there may be subtle cross-linguistic differences in the discourse-pragmatic interpretation of overt subject pronouns, with Italian being more restrictive than Iberian Spanish in the requirement that the antecedent for an overt subject pronoun be a non-subject referent (Filiaci, Sorace, & Carreiras, 2014). In this case, syntactic overlap is complete, but there is a partial mismatch in the way in which the same forms map onto different discourse functions cross-linguistically.

In other cases, generally when there is weaker typological similarity, it may be that some of the referential expressions that are available in the syntax of one language are not syntactically licensed in the other, or that different forms map onto the same discourse function. This is, for example, the case of Italian and English where null pronominal subjects and clitic pronominal objects are syntactically allowed in the former but not in the latter (Cardinaletti & Starke, 1999). In this scenario the same discourse function, e.g., reference to a discourse-prominent antecedent, would be mapped onto different forms: a null subject pronoun or a clitic object pronoun in Italian, and overt, non-clitic pronouns in the case of English.

The issue of partial structural overlap is one of the key aspects of the CLI hypothesis as originally formulated by Hulk and Müller (2000). It stands to reason that instances of CLI can only really be observed in cases in which the two languages differ, at least partially. If the overlap is complete both in the repertoire of linguistic forms and in the mapping to discourse functions, there is no obvious way to determine whether the form-function mapping in one language may be influencing the form-function mapping in the other language.

Inspired by Hulk and Müller’s proposal that phenomena at the syntax-pragmatics interface were those that would be more vulnerable to CLI, Paradis and Navarro (2003) and Serratrice, Sorace, and Paoli (2004) independently started a systematic investigation of the mapping of referential expressions as a function of discourse-pragmatic constraints. Although Hulk and Müller (2000) proposed
the interface between syntax and discourse-pragmatics as the likely locus of CLI, they did not actually investigate the discourse-pragmatic contexts of object omissions in their study. These two case studies of an English-Spanish child (Paradis & Navarro, 2003) and an English-Italian child (Serratrice et al., 2004) started a new line of inquiry. They made an initial contribution to our understanding of how children deal with the use of pronominal subjects when they are learning two languages where overt subject pronouns exist in both, but where they map onto different discourse functions.

Both Italian and Spanish are languages where null subjects are pragmatically optimal when they refer back to discourse-prominent subject antecedents (Carminati, 2002). As for overt subject pronouns, in Italian the preferred antecedent tends to be a non-subject non-discourse-prominent antecedent. The evidence for Spanish is somewhat less conclusive, partly because many different varieties of Spanish have been investigated in the literature (Morales, 1997), and partly because the methods vary across studies, ranging from naturalistic corpus data to online reading times. Evidence from corpus studies indicates that, similarly to Italian, overt subject pronouns in Spanish tend to signal a switch from the subject in the preceding tensed clause (Cameron, 1992; Cameron & Flores-Ferrán, 2004; Hochberg, 1986; Montrul, 2004; Morales, 1997; Otheguy, Zentella, & Livert, 2007; Silva-Corvalán 1994). More recently, however, reading time experiments showed that, at least for adult Spanish speakers, there was a smaller processing penalty in interpreting an overt subject pronoun as referring to a non-subject antecedent, than for adult Italian speakers (Filiaci et al., 2014). These findings apply to adult speakers and need to be replicated, but it is worth bearing in mind these subtle cross-linguistic differences when assessing the effect of typological differences in the language combinations of bilingual speakers.

Though there are some specific cross-linguistic differences between Italian and Spanish, these null-subject languages share the trait that the pragmatically optimal referential expression for a topical subject antecedent is a null subject. In English, the lack of syntactically licensed null subjects obliterates the discourse-pragmatic distinction observed in null-subject languages; the same overt pronominal subject forms map onto two distinct discourse-pragmatic functions when it comes to both topic maintenance and topic switch. Thus the English-Italian and English-Spanish combinations present a case of structural overlap (syntactically licensed overt subject pronouns) and a cross-linguistic discourse-pragmatic mismatch (overt subject pronouns in English > topic maintenance and topic switch discourse functions; overt subject pronouns in null-subject languages > only, or mainly, topic switch function). This presents an ideal situation in which to study how bilingual children deal with the form-function mapping in their two languages and the likelihood of CLI.
Chapter 13. Referential expressions in bilingual acquisition

The corpus evidence for CLI in the realization of subject arguments is based on a number of single case studies (Hacohen & Schaeffer, 2007; Haznedar, 2010; Juan-Garau & Pérez-Vidal, 2000; Paradis & Navarro, 2003; Serratrice et al., 2004), or studies including a small number of pre-school children recorded in interaction with their adult caregivers over a period of months or years (Hauser-Grüdl, Arencibia Guerra, Witzmann, Leray, & Müller, 2010; Liceras, Fernández-Fuertes, & Pérez-Tattam, 2008; Schmitz, Patuto, & Müller, 2012; Silva-Corvalán, 2014; Zwanziger, Allen, & Genesee, 2005). With the exception of the work by Hauser-Grüdl et al. and Schmitz et al., all of the studies mentioned above feature English as the non-null-subject language in the pair, and a variety of typologically different null-subject languages including Spanish, Turkish, Hebrew, Korean, Italian and Inuktitut. Although the common denominator across these corpus-based studies is the investigation of possible CLI at the level of sentential subjects, there is considerable variation in the way in which the issue is explored, and in the kind of aspects and determinants of CLI that are taken into consideration. The key questions focus around the likelihood of CLI, its directionality, the relative weight of children’s language competence and/or language exposure, and the importance of the quantity and quality of parental input. Unsurprisingly, not all the studies address all of these matters, and even when two or more studies do examine the same issue, the nature of the evidence brought to bear may be rather different.

5. How likely is CLI?

As more research is published on the subject of CLI, it is becoming increasingly clear that “the conditions on cross-linguistic influence – however they are formulated – are sufficient but not necessary” (Unsworth, 2013, p. 32). With respect to referential expressions, then, it is not necessarily the case that the null-subject / non-null-subject language combination will trigger CLI in the domain of subject realization, a fact that has started to emerge from the literature. Of the ten studies mentioned above, five did not find any evidence for CLI as measured by the relative proportion of null and overt subjects in the bilingual children’s production;
the bilingual children in these studies behaved in the monolingual range in terms of proportional subject realization (Juan-Garau & Pérez-Vidal, 2000; Liceras et al., 2008; Schmitz et al., 2012; Silva-Corvalán, 2014; Zwanziger et al., 2005). Failure to find an effect of CLI solely in terms of the proportion of overt and null subjects produced is however problematic if the assumption is that CLI is a syntax-interface phenomenon whose outcome is the lack of discourse-pragmatic appropriateness in the domain of referential expressions.

Crucially, with the exception of Silva-Corvalán (2014), none of the studies that failed to find an effect actually investigated whether there was a higher proportion of pragmatically inappropriate overt subjects in the null-subject language of the bilingual children. The nature of the data on which these calculations are carried out makes a comparison in purely numerical terms unsuitable to draw any firm conclusions as to the presence of CLI at the discourse-pragmatics interface. Unlike in experimental tasks, where the discourse contexts and the number of items are kept constant across participants, in the kind of adult-child naturalistic interactions that form the basis of these corpus studies, the type and number of discourse contexts inevitably vary. Without considering the discourse contexts in which the null and overt subjects appear we cannot a priori exclude that children might have omitted a subject when it was pragmatically required and, vice versa, used one when it was not. In the absence of a systematic investigation of the discourse-pragmatic context of subject realization, it is therefore unwarranted to rule out the possibility that (at least some) of the overt subjects produced by the bilingual children in their null-subject language may be in contexts in which they were not pragmatically optimal.

By the same token, among the studies listed above where the authors conclude that there is evidence for CLI in the domain of subject realization, two do not consider the discourse-pragmatic context at all (Hauser-Grüdl et al., 2010; Schmitz et al., 2012). Those that do (Hacohen & Schaeffer, 2007; Haznedar, 2010; Paradis & Navarro, 2003; Serratrice et al., 2004; Silva-Corvalán, 2014) consistently report the same finding: the bilingual children produced more discourse-pragmatically inappropriate overt subject pronouns in the null-subject language than the monolingual children; the proportion of null subjects in the non-null-subject language was unaffected.\(^3\) Across these studies, discourse-pragmatic inappropriateness is defined as the use of an overt pronominal form in low informativeness contexts, i.e., essentially contexts in which the referent is given and/or it is not in

\(^3\) In Silva-Corvalán’s (2014) study the picture is more complicated as the analysis covers a developmental period that is longer than in most of the other studies (1;6–5;11). Additionally, there are differences in terms of language exposure in the two children included in the study that, over time, correlate with different patterns of overt subject provision in Spanish.
need of disambiguation (see Allen et al., this volume, and Hickmann, Schimke, & Colonna, this volume, for an in-depth treatment of the relationship between informativeness and referential choice).

6. Rationale for the directionality of CLI

The results discussed in the previous section are consistent with the predictions concerning the directionality of CLI made by a number of authors. Paradis and Navarro (2003) formulated the hypothesis that the frequency of overt subject pronouns in the language in which they are syntactically obligatory (English) would lead the bilingual child to produce more of the same subjects in Spanish, the language in which they are obligatory in some, but not all, discourse contexts. Hacohen and Schaeffer’s (2007) take focused on the subset-superset relationship whereby the subset language (English) would affect the superset language (Hebrew) and result in an overproduction of pragmatically inappropriate subject pronouns. Haznedar (2007) did not explicitly formulate any predictions, but in her discussion she referred to the greater processing costs incurred by bilingual children in the integration of syntax and discourse-pragmatics in the null-subject language in the face of reduced input. What Haznedar did not add, but must be inferred from her argument, is that for the directionality to go from English to Turkish it has to be the case that the obligatoriness of subject pronouns in English plays a role in occasionally skewing the Turkish-English bilingual’s choices in favour of using a pronoun when one is not pragmatically optimal. Serratrice et al’s (2004) proposal was couched in terms of the relative markedness of overt subject pronouns in Italian – associated with [+topic] and [+focus] discourse features – and the bleaching of the discourse-pragmatic restrictions as an effect of regular use and exposure to English where the unmarked option is to use subject pronouns regardless of the value of topic and focus features. Finally, Silva-Corvalán (2014), capitalizing on a proposal by Serratrice (2007a), views CLI, and the overuse of pragmatically inappropriate overt subject pronouns in the Spanish of the two children in her case study, as the outcome of cross-linguistic priming. The argument is that the high-frequency [subject pronoun + verb] structure in English, in conjunction with the strengthening of English, as the language of greater exposure and use as the children get older, is responsible for the increase of overt subject pronouns in Spanish.

Although different authors have emphasized different representational and processing causes for the directionality of CLI, there is a general consensus that the dichotomy between the obligatoriness vs. the (pragmatically conditioned) optionality of overt pronominal subjects is key to the phenomenon. In addition to
focusing on this structural aspect, some have also questioned the role of input and children's language dominance and language exposure, although the relationship between directionality of CLI and language exposure has mostly been considered in experimental studies of older bilingual children (Serratrice, Sorace, Filiaci, & Baldo, 2012; Sorace, Serratrice, Filiaci, & Baldo, 2009).

7. The role of language exposure, language dominance and input quality in the use of referential expressions

The proposal by Silva-Corvalán (2014) explicitly factors in exposure and use of English as an external determinant of CLI in the referential domain. This suggests that some form of dominance – whether it be conceived in terms of the child's own linguistic competence, or of the amount of exposure, or both – is relevant to CLI. We will return to this point below, but first we turn to the issue of the quality of the input to bilingual children, and whether any differences either in terms of non-nativeness or language attrition may be considered as underlying causes of CLI.

Paradis and Navarro (2003) explicitly hypothesized that the higher than expected proportion of low informativeness overt pronominal subjects in the Spanish of the English-Spanish bilingual child in their study could be ascribed to the proportion of low informativeness overt pronominal subjects in the non-native Spanish input from her mother, and in the Panama Spanish variety spoken by her father. Paradis and Navarro (2003) made an important point that has not received much attention since: the nature of bilingual families is such that children are unlikely to be exposed to the same quality of input as monolingual children, in that both non-native and attrited speech are likely to be a feature of their language environment. With specific reference to the distribution of pronominal subjects in null-subject languages, we know that, similarly to bilingual children, both second language learners (Montrul, 2004) and attrited adult L1 speakers (Tsimpli, Sorace, Heycock, & Filiaci, 2004) have issues with the discourse-pragmatic constraints on overt pronominal subjects. It is, therefore, a limitation of the studies that have appeared on the issue of CLI and pronominal subjects that no attention has specifically been paid to the nature of the input in bilingual families where a non-trivial amount of input may be non-native or attrited.

The only study, among the ones being reviewed here, that directly addressed the question of the quality of the input in relation to subject realization is Hauser-Grüdl et al. (2010). The authors took exception to Paradis and Navarro's (2003) proposal that the overproduction of pragmatically inappropriate overt subject pronouns in Spanish is the result of contact-modified input. To prove their point, they showed that the Germany-based Italian-speaking parents of five Italian-
German bilingual children who did overproduce overt subjects in Italian, did not in fact use more overt subjects than adult monolingual speakers living in Italy. Their argument was that CLI in the domain of subjects cannot therefore be the result of exposure to attrited parental speech. Although, of course, nobody would dispute that contact-modified input is but one of the cause of CLI, there are several problems with Hauser-Grüdl et al.’s (2010) argument. Firstly, they conflated both pronominal and lexical subjects in their count when we know that the issue is really with the use of pronouns, rather than with subjects tout court. Secondly, as suggested earlier, in the absence of a systematic discourse-pragmatic analysis it is not possible to conclude whether the null, pronominal and lexical subjects that were included in the count also appeared in target-like discourse-pragmatic contexts. Thirdly, the lack of a positive association between the proportion of null and overt subjects between parental and child speech only shows that contact-modified speech is not necessary for CLI to occur, but it cannot rule it out. Although Hauser-Grüdl et al.’s (2010) arguments are far from conclusive they do address the important question of the role of attrited and/or non-native input in bilingual acquisition – an issue that has unfortunately largely been neglected in the investigation of the possible causes of CLI.

In contrast with the paucity of data on the quality of the input in bilingual acquisition, more research has been devoted to the relationship between input quantity and language dominance and CLI. If language exposure is a relatively straightforward concept to define and measure, typically counting the number of hours the child is exposed to her two languages over the course of a specified time period, the notion of language dominance is less so. There is no general agreement on which measures best express language dominance; indicators used include Mean Length of Utterance (MLU), Upper Bound (i.e., the longest utterance in a given transcript), the number of (multimorphemic) utterances, vocabulary size, and number of words per minute. Amount of exposure has also on occasion been used as a proxy for language dominance. In general terms, the frequency of use of a language and the amount of input received have repeatedly been found to be good predictors of children's language proficiency (Gathercole & Thomas, 2009; Gutiérrez-Clellen & Kreiter, 2003; Hoff, Core, Place, Ruminche, Señor, & Parra, 2012).

With specific reference to CLI in the domain of subject realization, most of the studies reviewed here offer some general background information on the amount and pattern of exposure to the two languages, although, with the exception of Silva-Corvalán (2014), there has so far been no explicit attempt in corpus studies to relate the children’s overproduction of pragmatically inappropriate overt subjects to diminished exposure to the null-subject language. Hauser-Grüdl et al. (2010), on the other hand, are the only ones who postulate a positive relationship
between the likelihood of CLI in the domain of subject realization and language dominance defined by fluency (i.e., number of words per minute). Previous work on bilingual children has found a positive association between fluency (number of words per minute) and MLU, a traditional measure of language competence (Arencibia Guerra, 2008). Hauser-Grüdl et al.’s argument is that fluency should be included as an explanatory variable, alongside structural complexity, in accounting for the directionality of CLI and for the degree of individual variation that is observed across different children.

The formulation of a multifactorial model is bound to be more likely than a simple factorial model to account for CLI. CLI is a complex phenomenon where structural overlap is a necessary condition, but it is not sufficient to account for the ways in which child-internal (age, dominance, speed of processing) and child-external factors (language environment; language exposure) interact. Corpus studies investigating CLI in pre-school children in the domain of subject realization have highlighted a number of these factors, and some of them have attempted to consider more than a purely structural explanation for the phenomenon. This approach has been further elaborated by a handful of studies investigating the issue in older school-age children.

8. Older bilingual children: Experimental evidence for the comprehension and production of referential expressions

Research investigating CLI in the domain of subject pronouns in pre-school children has so far relied on corpus data of naturalistic adult-child interaction. A few studies on older school-age children have instead used comprehension and production experiments, as well as narrative data from larger groups of children, to explore this issue (Argyri & Sorace, 2007; Serratrice, 2007a, 2007b; Sorace et al., 2009). The move from single case studies, or case studies of a small number of children, has allowed researchers to abstract away from individual variation and to shed some light on general tendencies in the bilingual population. Moreover, comparisons between different language combinations (e.g., Italian-English vs. Italian-Spanish), a cross-sectional approach including younger and older children, and the inclusion of groups of children with the same language combination but with different majority community languages (e.g., Greek-English bilinguals living in the UK or in Greece) have all contributed to a better understanding of the determinants of CLI in older bilinguals.

When comparing the same phenomena across different studies, there are both similarities and differences. In terms of the acceptability of null and overt subject pronouns in topic maintenance contexts in null-subject languages like
Italian and Greek, the evidence for the unidirectionality of CLI is rather unambiguous. Regardless of dominance or language of the community, the bilingual children's English is not affected and ungrammatical null subjects are rejected as such in forced choice tasks (Argyri & Sorace, 2007; Sorace et al., 2009).

As for the acceptability of overt pronouns in topic maintenance contexts, the evidence is somewhat less clear-cut. In the results of a picture verification task, Serratrice (2007a) reported that 8-year-old Italian-English bilingual children living in Italy accepted significantly more pragmatically inappropriate anaphoric overt pronouns in topic maintenance contexts than both monolingual peers and adults. There was, however, no significant difference in the rate of acceptance of pragmatically sub-optimal cataphoric overt pronouns between the bilingual and the monolingual children, although both groups accepted significantly more than the adult monolinguals. In the findings of a forced-choice acceptability task, Argyri and Sorace (2007) reported that English-Greek bilingual children living in the UK, who had English as their dominant language, accepted more pragmatically inappropriate overt pronouns in Greek than monolingual adults. They were, however, not significantly different from a group of Greek-English bilinguals who were living in Greece and were Greek-dominant, or from a group of monolingual Greek children. In an elicited production task with the same children, Argyri and Sorace found that all groups, regardless of age or of bilingual status, behaved at ceiling and used only null subjects in topic maintenance contexts.

Finally Sorace et al. (2009) compared English-Italian younger (6- to 7-year-old) and older (8- to 10-year-old) bilinguals living in Italy or in the UK with monolingual peers in English and Italian, and with Spanish-Italian children living in Spain. With respect to the acceptability of pragmatically inappropriate overt subject pronouns in Italian, Sorace et al. (2009) uncovered developmental differences. In the younger group, only the UK-based bilinguals chose significantly more overt pronouns in topic maintenance contexts than their monolingual or bilingual peers; in the older age range, it was the Spanish-Italian group that opted significantly more frequently for non-target overt subject pronouns. This latter finding was interpreted by Sorace et al. (2009) as an indication that sub-optimal discourse-pragmatic choices may not necessarily be exclusively the outcome of CLI, but may be a result of bilingualism per se. In essence, a bilingual speaker’s resources may be particularly taxed in cases where the choice of a linguistic form requires the subtle coordination of syntax and discourse pragmatics, as in the case of referential expressions. On the assumption that bilingual speakers can never completely inhibit the language not in use, a degree of competition across the two languages is to be expected when it comes to choosing between different referential forms, and this may result in non-target selection. In cases like the one of pronominal choice, where speakers have a number of referential expressions
that are more or less pragmatically acceptable in a given context, the partial mismatch in mapping between form and discourse function in two languages may occasionally result in the selection of a form that is less than ideal in one of the two languages.

Although this interpretation of the data is a plausible one, and the pairing Spanish-Italian was specifically selected for typological comparison purposes with English-Italian, recent findings on the online behavior of monolingual Spanish speakers (Filiaci et al., 2014) have cast some doubt on the hypothesis that the non-target behaviour of the older Spanish-Italian bilinguals is due to the processing overload of dealing with two languages. As mentioned earlier, Filiaci et al.’s (2014) results need to be replicated, and at the moment they are only available for adult monolingual Spanish speakers. Nonetheless, it appears that Iberian Spanish speakers may be more tolerant than Italian speakers in allowing overt subject pronouns in topic maintenance contexts. If this is the case, then there are implications for Sorace et al.’s (2009) Spanish findings. If the Spanish-Italian combination is more similar to the English-Italian combination than previously thought, then the behavior of the Spanish-Italian children in the study may be more of a result of CLI than bilingualism per se. Only further empirical evidence will tease apart these two options; for the moment the possibility remains that neither CLI, nor the processing overload associated with dealing with two languages, can be ruled out as potentially valid explanations.

In addition to the null/overt realization of subject pronouns, Argyri and Sorace (2007) also investigated whether CLI would affect the interaction between focus and word order. In Greek, non-contrastive wide-focus contexts require the placement of the subject after the verb (Alexopoulou, 1999), while in English this syntactic position is not available. The prediction in the study was that the obligatoriness of the preverbal position in English would lead to a significantly higher acceptance and use of preverbal subjects in Greek, regardless of focus context, for the English-Greek bilinguals compared to their monolingual peers. The results in both the elicited production task and in the forced-choice acceptability task did indeed show a unidirectional influence from English to Greek, but only in the case of the English-dominant bilinguals, who produced and accepted significantly more preverbal subjects than both Greek-dominant bilinguals and monolingual children and adults.

From the mixed picture arising from studies of older bilingual children in the domain of pronominal referential choice, it is becoming clear that the simultaneous acquisition of two languages does not necessarily trigger CLI. Rather, the relative exposure to the non-null-subject language plays an important role. Another key issue is that, when it comes to subtle referential choices that are constrained by discourse-pragmatics, we need to factor in development effects alongside
bilingualism. Some of the studies reviewed here reported that both monolingual and bilingual children were significantly less pragmatically appropriate than monolingual adults in their choices. It may therefore be relevant to think about bilingual children’s sub-optimal selections as a result of delay that is driven, at least partially, by reduced language exposure.

The next section will turn to an overview of the research on object pronouns in bilingual children and to evidence for CLI, for its directionality, for the role of input, and for the role of a bilingual effect independent of CLI.

9. CLI and direct object pronouns

The CLI hypothesis was originally formulated in connection with the higher than expected omission of direct objects in the Romance language of bilingual children exposed to a topic-drop Germanic language (Hulk & Müller, 2000). However, despite the original focus on pronominal objects, there is currently a less extensive literature on object than on subject arguments. Similarly to the case of pronominal subjects, research on the realization/omission of pronominal objects in childhood bilingualism has mostly been investigated in language pairs in which the null option is licensed by only one of the two languages – either topic-drop languages like German and Dutch (e.g., Hulk & Müller, 2000; Müller & Hulk, 2001) or null argument languages like Cantonese (e.g., Yip & Matthews, 2000). The other language of the pair is a language where null objects are not syntactically licensed (e.g., English, French, Italian). In the case of French and Italian, two SVO Romance languages, the canonical post-verbal object position is empty when object arguments are pronominalized as they are realized as preverbal clitics (e.g., Laura mangia la mela > Laura lai mangia ti ‘Laura eats the apple > Laura eats it’). The hypothesis is that the pervasiveness of null topicalized objects in German or Dutch (e.g., ec, Hab’ ich t, schon gesehen ‘I have already seen (him)’) will reinforce the availability of a null post-verbal position in Romance languages and therefore lead to more null objects, and for a longer period of time, than is typically observed in monolingual Romance acquisition.

The common finding is that bilingual exposure to a language where null objects are allowed leads to unidirectional CLI where a higher object omission rate is observed in the language where null objects are typically not allowed. This was the case of Italian and French in the Hulk and Müller studies, and of English in a case of Cantonese-English bilingualism (Yip & Matthews, 2000). In cases in which neither language licenses null objects (e.g., English-Italian, English-French, English-Greek) there should, in principle, be no differences between bilingual and monolingual acquisition as there is no case to be made for CLI. This is what
Serratrice et al. (2004) found for one English-Italian bilingual child between the ages of 1;10 and 4;6; provision of object arguments in the child’s corpus was no different from that of monolingual children in either language. Argyri and Sorace (2007) reached a similar conclusion in their experimental study of Greek-English bilingual 8-year-olds. They tested bilinguals’ and monolinguals’ knowledge of the distribution of preverbal object clitics in Greek in both an acceptability judgment task and an elicited production task, finding no differences as a function of either language background (monolingual vs. bilingual) or language of the community (Greek vs. English). Children’s performance was at ceiling in producing and accepting object clitics in preverbal position, and in rejecting ungrammatical postverbal object clitics in Greek, and, conversely, in only accepting and producing postverbal strong pronouns in English.

Serratrice et al. (2012) also tested the intuitions of younger (6- to 7-year-old) and older (8- to 10-year-old) English-Italian and Spanish-Italian bilinguals on the grammaticality of object pronouns in both languages in both [-focus] and [+focus] conditions. Unlike Argyri and Sorace (2007), Serratrice et al. (2012) did find differences in the judgments of bilingual and monolingual children, alongside some similarities. Like the Greek-English bilinguals, regardless of age and language of the community (English, Italian), the English-Italian bilinguals performed at ceiling in rejecting ungrammatical sentences in the [-focus] condition in which the object pronoun was in preverbal position in English (e.g., *He him saw). In [+focus] conditions, where participants had to judge between a pragmatically appropriate prosodically focused pronoun and a pragmatically sub-optimal non-focused pronoun (e.g., Did Mickey see Donald or Minnie? He saw HER vs. He saw her), there were no significant effects of language background (monolingual vs. bilingual) or language of the community (English vs. Italian). There was an effect of age, however, with all children improving from chance performance in the younger group to around 60% accuracy in the older group. These findings suggest that coordinating prosodic, syntactic and pragmatic information is a demanding task and that, in addition to the effect of bilingualism, we need to factor in a developmental effect. Moreover, failure to find acceptance for ungrammatical constructions with preverbal strong pronouns suggests that “cross-linguistic influence must obey a principle of isomorphism, and that it can only take place when there is morphosyntactic equivalence between constructions across the two languages” (Serratrice et al., 2012, p. 731). In the case of strong pronouns in English and clitic pronouns in Italian there is no morphosyntactic equivalence and the prediction is therefore that CLI should not take place. As for the Italian data, the hypothesis was that the bilingual children would be more likely to accept postverbal strong pronouns as pragmatically acceptable in a [-focus] context given that this word order is also attested in Italian – but only in [+focus]
contexts – and that it is the only available word order for English (strong) pronouns. The results of the acceptability judgment showed a trend whereby language of the community made a difference. The younger bilinguals in the UK accepted pragmatically inappropriate postverbal strong pronouns in [–focus] contexts (e.g., "Vede lei’ (He/she) sees her’) 35% of the time compared to 11–23% for the other groups of bilingual and monolingual children. The same pattern held for the older bilinguals, although the gap was narrower with the bilinguals in the UK accepting inappropriate pronouns at the rate of 20% compared to 9–10% in the other bilingual and monolingual groups. As for the [+focus] condition in Italian, where the pragmatically appropriate option is a prosodically marked strong pronoun in post-verbal position, all groups fared rather poorly with correct acceptance and rejection rates below chance. The only significant trend found was for the monolingual Italian-speaking children who reached 70% accuracy rates in the older group. This pattern of results suggests that, for older children who have moved away from an initial object-drop phase, CLI in the domain of pronominal objects is still possible when it comes to deciding which word order (pre- or post-verbal) is associated with [+– focus] contexts. The less than target-like performance of the monolingual children in the Serratrice et al. (2012) experiments also confirms that, whatever effect we can ascribe to bilingualism in accounting for children's non-target performance, we still need to factor in age effects to explain a greater amount of variance. Language exposure, as measured by language of the community, also played an important role, with children living in the UK showing more pronounced effects of unidirectional influence from English to Italian.

Taking a developmental step backwards, we now turn to earlier phases in which both monolingual and bilingual children omit object arguments, regardless of whether they are acquiring a null-object or a null-topic language, and assess the extent to which bilingualism per se, rather than partial structural overlap in a specific language combination, results in higher than expected omission rates in the domain of pronominal objects. The hypothesis of a bilingual effect has been formulated over the last few years in a number of studies on pronominal object omission by Pérez-Leroux, Pirvulescu, Roberge and colleagues (Pérez-Leroux, Pirvulescu, & Roberge, 2009; Pirvulescu, Pérez-Leroux, & Roberge, 2012; Pirvulescu, Pérez-Leroux, Roberge, Strik, & Thomas, 2014). The underlying rationale for the argument is that there is considerable cross-linguistic variation as to whether null objects are allowed and how they are licensed – from discourse-linked referential null objects (e.g., Cantonese, Mandarin, Japanese), to null objects linked to topicalized antecedents (e.g., German, Dutch), to deictic null objects (e.g., English, Italian), or to null objects licensed by individuated contexts where an antecedent is present in the preceding discourse (e.g., English, French, Italian). Given this state of affairs, Pérez-Leroux and colleagues capitalize on the
proposal that the default setting made available by Universal Grammar for direct objects is that of a bundle of features that are active but may or may not be silent, depending on the language. Pérez-Leroux, Pirvulescu and Roberge (2008) make the additional stipulation that both unergative and transitive verbs allow for a structural direct object position, regardless of whether the direct object is realized overtly, and that child grammar starts out with this default option. Depending on which language(s) children are exposed to, the relative availability of null objects will sit alongside their overt realization. The extent to which children converge on the adult target will depend on how unambiguous the input is. In languages where null objects are allowed, at least in certain discourse-licensed contexts, it will take children longer to figure out pronominal object realization. In cases of bilingual acquisition, input ambiguity is compounded by a reduced amount of input in each language and will result in “longer retention of the default representation in both languages (i.e., object omission)” (Pirvulescu et al., 2014, p. 501).

In two earlier experimental studies with French-English bilingual pre-school children (Pérez-Leroux et al., 2009; Pirvulescu et al., 2012), higher and protracted rates of pronominal object omission were reported in the bilingual children’s French when compared to monolingual French peers. The main limitation of both studies was the investigation of only the French output, and it was therefore not clear to what extent English was responsible for the observed pattern in French. Pirvulescu et al. (2014) addressed the issue by testing bilingual French-English children between the ages of 3 and 6 with different patterns of language dominance as measured by a parental questionnaire on language exposure (balanced bilinguals, French-dominant and English-dominant). The bilingual children took part in a production task eliciting direct object pronouns in both languages, and were compared to English-speaking and French-speaking monolingual peers. The results show a previously attested cross-linguistic asymmetry with more object omissions in French than in English for both monolinguals and bilinguals. More importantly, the bilinguals omitted more objects than the monolingual children in both languages, and there was a significant effect for language dominance in French whereby the balanced bilinguals omitted more objects than the children who were either dominant in French or in English. Interestingly, there was no significant effect of the amount of English input: the children who had less exposure in French did not omit more objects than either the balanced bilinguals or those who were dominant in French. The conclusion reached by Pirvulescu et al. (2014) was that bilingualism per se may lead to delays in contexts where there is a structural default, and where the variability of the input in the target grammar, combined with reduced amount of exposure, decreases the relevant data required to settle on the target grammar.
The literature on referential choice in the domain of object pronouns paints a complex developmental picture. Similarly to what was observed for subject pronouns, the realization and distribution of object pronouns in bilingual contexts is affected, to different extents, by language exposure and by the morpho-syntax of the language pair. With specific reference to the cross-linguistic differences existing in the availability of object clitics it appears that, while clitics are vulnerable in terms of omission in the sense that bilingual children may persevere for longer in an early null-stage object phase, they are resilient to word order errors. As argued by Serratrice et al. (2012), bilingual children are sensitive to the different morpho-syntactic status of clitics vs. strong pronouns and do not accept clitics in ungrammatical post-verbal positions.

10. Conclusion

This overview of the literature on the production and comprehension of pronominal referential expressions has highlighted the complex nature of a linguistic phenomenon that requires the coordination of sophisticated knowledge in not one but two languages. Some of the challenges faced by bilingual children are the same as those encountered by monolingual peers: creating discourse-relevant form-function mappings that take into account a set of morphological, syntactic, prosodic and pragmatic requirements. Some of these are language-universal (e.g., the use of more informative NPs for less accessible referents) while others are language-specific (e.g., the presence of clitic pronouns), and language-relevant input in each language will determine children’s language-specific choices. How similar or different bilingual children’s choices are from monolingual peers will depend on a range of factors including the language combination they are acquiring and hence the degree of structural overlap across the two languages, the degree of variability in the target language(s), and the amount of input that is necessary to home in on the target grammar(s).

Corpus and experimental data have provided behavioural evidence of what bilingual children can do, i.e., about the end output. What we still need to know more about is how bilingual children differ, or not, from monolingual children in the process that leads to these behavioural choices, and what other non-linguistic predictors can tell us about both the process and the final outcome. Online studies of bilingual children’s referential choices (see Sekerina’s overview in this volume for monolingual children) and the investigation of cognitive correlates of referential choice (see De Cat’s overview in this volume) will add two other important pieces to this complex puzzle.
Acknowledgments

Many thanks to Carmen Silva-Corvalán and Shanley Allen for very valuable feedback on earlier versions of this chapter. All errors, inaccuracies and infelicitous interpretations of the literature are, of course, our own.

References


Cameron, R., & Flores-Ferrán, N. (2004). Perseveration of subject expression across regional dialects of Spanish. Spanish in Context, 1(1), 41–65. DOI: 10.1075/sic.1.1.05cam


Chapter 13. Referential expressions in bilingual acquisition


