Comparing anaphora resolution in early and late Brazilian Portuguese-European Portuguese bidialectal bilinguals

Article

Accepted Version


It is advisable to refer to the publisher’s version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1075/resla.29.2.03cas

Publisher: John Benjamins Publishing Company

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur
CentAUR
Central Archive at the University of Reading
Reading’s research outputs online
Comparing anaphora resolution in early and late Brazilian Portuguese-European Portuguese bidialectal bilinguals

Tammer Castro¹, Jason Rothman¹² and Marit Westergaard¹
¹CASTL – UiT Arctic University of Norway / ²University of Reading

The present study examines anaphora resolution in two groups of speakers exposed to Brazilian and European Portuguese (BP and EP, respectively), considering the different null subject distribution in these languages. Our research question is whether late BP-EP bilinguals (age of EP onset: 29.1) and heritage BP speakers raised in Portugal (age of EP onset 5.6), tested in both dialects, will pattern like the native controls or display some effects of EP in their native BP or vice-versa. This is an interesting question in light of the Interface Hypothesis, which claims that external interfaces should be subject to general bilingualism effects irrespective of language pairing and age (Sorace, 2011). The results show that age has an effect, as the heritage speakers do not perform like the late learners, and that the high degree of typological proximity between the two languages rules out the possibility of late bidialectal acquisition.

Keywords: null subjects, Portuguese, bilingualism, bidialectalism, attrition

1. Introduction

Over the past few decades, empirical studies on adult second language (L2) and heritage language (HL) bilingualism have made valuable contributions to our understanding of bilingual language acquisition and processing (e.g., Sorace & Filiaci, 2006; Montrul, 2008; Rothman, 2009; Schmid & Hopp, 2014). While HL and L2 acquisition are generally analyzed separately, they can be compared to one another to tackle fundamental theoretical debates within the field, considering multiple variables at once (see Montrul, 2008, for a review). The present study targets heritage Brazilian Portuguese (BP) speakers who have grown up in Portugal and Brazilians who are adult L2 learners of European Portuguese (EP), with the purpose of shedding light on both of these subfields individually and where they
interact. Since heritage speakers share features with both first and second language acquirers, in this study we compare them to both groups in order to better understand the similarities and differences between them, hoping to trigger future discussions in the general field of language acquisition.

It is generally the case that studies on L2 acquisition and HL bilingualism analyze language pairings involving typologically distinct languages such as Russian-English or Spanish-English (e.g., Ionin, Zubizarreta, & Maldonado, 2008) or German-EP (e.g., Rinke & Flores, 2014) but a few have also tested closely related languages (e.g., Bini, 1993; Filiaci, 2010, for Spanish-Italian; Montrul, Dias, & Santos, 2011, for Spanish-BP). The present study goes even further by investigating both L2 and HL acquisition in the context of bidialectalism. BP and EP have been described in the literature as displaying a high degree of difference at all linguistic levels (e.g., lexis, phonology, syntax, morpho-syntax) (see Costa, Lobo, & Silva, 2009; Barbosa, Duarte, & Kato, 2005, for discussion). Even though BP and EP share a very high degree of mutual intelligibility, both varieties display distinctive features that have led scholars to consider them separate languages (Perini, 2010; de Castilho, 2010). This study does not deal with a case of naturalistic acquisition of two variants with shared syntax (e.g. Shockey, 1984, for American and British English), but rather we target morpho-syntactic differences between two mutually intelligible Portuguese variants. As a consequence of typological similarity, researchers have shown that negative transfer is likely to occur to a much larger extent between two closely related dialects than between two languages that are typologically different (see Wolfram & Schilling-Estes, 1998; Long, 2007). Thus, the overlap between the two variants could potentially hinder the acquisition of these morpho-syntactic distinctions.

This study explores a domain related to referential subjects. The distribution of null/overt subjects across EP displays the typical Romance pattern, which allows root null subjects, referential embedded null subjects and null expletives (Costa & Pratas, 2015). BP, on the other hand, shows a mixed pattern which includes many more overt subjects than what is generally seen in Romance null-subject languages (NSLs) (e.g., Duarte, 1993, 1995; Barbosa et al., 2005). We test two sets of EP-BP bilinguals: (i) adult heritage speakers (HS) of BP who learned EP as children and (ii) adult BP natives who learned EP in Portugal in adulthood. We investigate their performance in a comprehension task on matrix subject vs. object orientation for anaphoric resolution of embedded null vs. overt subject pronouns. Both groups are tested in BP and EP, which ensures a unique comparative data set that sheds light on both HL and L2 acquisition. We compare HSs and L2ers precisely because these two populations differ with respect to crucial factors such as age of L2 acquisition and amount of first and second language use, besides the degree of first language (L1) proficiency and maturational state (Montrul, 2008). We take
Comparing anaphora resolution

into consideration the syntactic differences between BP and EP to test the extent to which: (a) the target groups understand these differences and acquire both patterns, (b) age determines directionality of influence between BP and EP, and (c) the high level of typological proximity between BP and EP in a bidialectal bilingual scenario influences the answers to (a) and (b) as compared to what previous research on HL and L2 acquisition has shown.

In this study, we focus on how the different distribution of null subjects in BP and EP plays a role in ambiguous forward anaphora contexts. In light of the Position of Antecedent Hypothesis (PAH) (Carminati, 2002, 2005), the contrast between a null and an overt pronoun in an embedded context yields different co-reference preferences. We will observe the participants’ understanding of how the inclusion of an overt pronoun generally triggers co-reference in lower positions in the phrase structure, whereas a null embedded subject establishes co-reference with the constituent in Spec-IP: the subject of the matrix clause. As shown in Sorace and Filiaci (2006), the PAH does not belong strictly to syntax, but involves the interface between syntax and discourse pragmatics. Previous studies testing the PAH have focused exclusively on contrasting either NSLs with non-null-subject languages (NNSLs) (e.g. Sorace & Filiaci, 2006; Keating, Jegerski, & VanPatten, 2016, among others) or two NSLs such as Italian and Spanish (e.g., Filiaci, 2010; Filiaci, Sorace, & Carreiras, 2014). The present study is innovative in that it contrasts EP, a typical NSL, and BP, a language that, while not a NNSL per se, has been described in the literature as a partial-NSL because of its mixed pattern (Holmberg, Nayudu, & Sheehan, 2009; Holmberg & Sheehan, 2010). BP has gone through a process of change regarding subject pro-drop, more specifically in 3rd person contexts. Duarte (1993, 1995) claims that BP is moving from being a NSL to becoming a NNSL, for reasons linked to the gradual simplification of its pronominal paradigm in the past century. While EP speakers make use of overt embedded subjects co-referential with matrix subjects only when indicating stress, emphasis or contrast, BP speakers allow for this co-reference in neutral situations (Ferreira, 2009). Thus, the acquisition of both BP and EP within our target populations necessarily involves acquiring different representations at the syntax-discourse interface.

Given the differences between the distribution of null subjects in both varieties, we test whether the two target groups: (a) understand the syntactic differences between the two varieties and are able to use subjects in appropriate contexts in both BP and EP-modes, (b) whether there is cross-linguistic influence or possible attrition in either direction such that the BP and/or EP grammars of these bilinguals are different from monolingual baselines, and (c) whether cross-linguistic influence, if evident, is conditioned by degree of dominance and/or age of arrival in Portugal. As contextualized above, we are especially interested in what
this language pairing and these data can add to theories of adult L2 acquisition and debates within HL acquisition.

2. Background

2.1 Adult second language and heritage language acquisition

Studies on adult L2 acquisition have generally converged on how this type of acquisition differs from native monolingual acquisition with respect to development and ultimate attainment (see Meisel, 2011; Slabakova, 2013). Factors such as the quality of the input and the conditions of acquisition (naturalistic vs. instructional) have been shown to play a role in the variability of L2 learners, though this variability is also found at the intra-speaker level (Bullock & Toribio, 2006). What remains unclear is what the differences between monolinguals and L2 learners indicate regarding successful acquisition (Slabakova, 2009, 2013). While ultimate attainment is shown to be possible when L2 acquisition begins in early childhood (Abrahamsson & Hyltenstam, 2009), this is generally not the case in adult L2 acquisition (see Meisel, 2009, 2011, for discussion).

Studies on HL acquisition have also been a growing trend in recent years (see Benmamoun, Montrul, & Polinsky, 2013, for a review). HS bilinguals, like L1 acquirers, first have exposure to the heritage language at home, via naturalistic input. The HL is, however, not the majority language of the larger society. Montrul (2011, p. 157) uses the following categorization for different types of HSs:

(a) simultaneous bilinguals, those exposed to the heritage and the majority language before the age of 3–4;

(b) sequential bilinguals or child L2 learners, those exposed to the heritage language at home until age 4–5 and to the majority language once they start preschool;

(c) late child L2 learners, children monolingual in the heritage language, who received some elementary schooling in their home country and immigrated around ages 7–8

The reasons for this subdivision according to age is linked to different processing strategies depending on the age of onset (see Meisel, 2011, for child L2 learning). Though typically dominant in the majority language, all HSs are native speakers of the HL, and they grow up in a bilingual setting with various degrees of dominance of the HL. Some learners manage to reach very high levels of proficiency and turn out to be balanced bilinguals, whereas others are significantly influenced by the
majority language and lag behind in the development of their L1 (see Rothman & Treffers-Daller, 2014). Many authors have attempted to explain this variability (e.g., Rothman, 2007; Putnam & Sánchez, 2013), which has been shown to be linked to the quality of input and amount of exposure to the HL, possible attrition of acquired structures (the features are acquired and subsequently lost), and even incomplete acquisition of structures (Montrul, 2008). There seems to be a general consensus that these are not mutually exclusive, and can all explain variability in HL acquisition to a certain extent.

The questions regarding adult L2 and HL acquisition typically target different variables, largely due to the crucial distinctions between both sets of learners. While HSs acquire the HL in a naturalistic setting early on, adult L2 learners are only exposed to the L2 after puberty, and typically in a classroom setting. As mentioned above, however, both types of learners also display a distinct acquisition pattern from that of L1 acquirers. The present study therefore takes a novel approach by directly comparing L2 and HL acquisition.

2.2 The Position of Antecedent Hypothesis

The Position of Antecedent Hypothesis (PAH) was first proposed by Carminati (2002) in a study investigating the processing of Italian null and overt subject pronouns in intra-sentential anaphora. According to the PAH, null pronouns in embedded contexts have a stronger tendency to be linked to the constituent in Spec IP, typically the subject of the matrix clause. Conversely, overt pronouns prefer to have antecedents in a lower syntactic position. This seems to apply for both canonical and non-canonical subjects (as shown in sentence (1) with a dative subject) as well as referential and non-referential subjects, as illustrated in (2a) and (2b), respectively:

(1) Poichè a Giovanni non piace affatto Enzo, Øi cerca di evitarlo.  
    ‘Because to G. does not please E. at all, Ø tries to avoid him.’  
    (Carminati, 2002, p. 104)

(2) a. Al colloquio per il posto di assistente di volo, ogni candidatai ha detto 
    che Øi/lei vorrebbe prendere le ferie ad agosto.  
    ‘At the interview for the post of air steward, every candidate (fem) has 
    said that Ø/she would like to have (her) vacation in August.’

    b. Al colloquio per il posto di assistente di volo, ognunoi ha detto che Øi/
    lui vorrebbe prendere le ferie ad agosto.  
    ‘At the interview for the post of air steward, everyone has said that Ø/he 
    would like to take (his) vacation in August.’  
    (Carminati, 2002, p. 283)
Further to these examples in Italian, the PAH was also shown to hold for Spanish (Bel & García-Alcaraz, 2015) and Romanian (Geber, 2006), presumably due to the fact that these languages have null subject distributions that are very similar to Italian, i.e., typical NSLs. It must be noted that in these languages, while the co-reference preferences for null pronouns are generally limited to Spec IP, overt pronouns are less strict, allowing for possible co-reference higher in the phrase structure. As pointed out by Carminati (2002), the PAH is more strict in ambiguous contexts, as in (3a), than when there is no ambiguity, as shown in (3b):

(3) a. Maria scrieveva frequentemente a Piera quando lei era negli Stati Uniti. ‘Maria wrote frequently to Piera when she was in the USA.’ (Carminati, 2002, p. 78)

b. Quando Maria ha chiamato Mario, Ø era contento. ‘When Maria has called Mario, Ø was (3rd person sg) happy.’ (Carminati, 2002, p. 187)

Considering the fact that most PAH violations are not ungrammatical, but rather inappropriate, Sorace and Filiaci (2006) argue that these contexts are necessarily at the interface between syntax and discourse pragmatics, as they involve the syntactic conditions on the licensing of null subjects and the discourse conditions that contextualize the distribution of the pronominal forms. In their study, they tested whether L1 English-L2 Italian (near natives) acquired the syntactic and pragmatic conditions on the licensing of null subjects in Italian, which is necessary to establish target co-reference in forward and backward anaphora. Given the crucial difference between English (NNSL) and Italian (NSL), the authors predicted that more differences between natives and bilinguals would be found in contexts with overt pronouns than those with null pronouns, which turned out to be the case. This is due to the ambiguity generated by the presence of the overt pronoun in certain contexts. Given that overt embedded pronouns can have different referents, whereas null pronouns in that context typically trigger one specific referent, the authors show that it becomes easier for both natives and bilinguals to establish target co-reference with null pronouns than with overt pronouns.

2.3 The Interface Hypothesis

In order to account for non-convergence and optionality at advanced stages of adult L2 acquisition, Sorace and colleagues developed the Interface Hypothesis (IH) (Sorace, 2000, 2003; Tsimpli, Sorace, Heycock, & Filiaci, 2004; Sorace & Filiaci, 2006), with the goal of unifying findings from bilingual L1 acquisition, L1 attrition and adult L2 acquisition. The original idea behind the IH was that L2 acquisition was more likely to be successful in contexts involving purely syntactic
comparisons than in contexts where the syntax interfaces with interpretable domains, such as discourse/pragmatics or semantics. The IH also predicted a similar pattern for bilingual L1 acquisition and early stages of L1 attrition concerning the same structures (see Sorace, 2000; Tsimpli et al., 2004). This pattern consists of residual optionality in later stages of L2 development, and emerging optionality in first-generation speakers as a result of language contact. This optionality, whether residual (in L2 acquisition) or emerging (in L1 attrition), is related to the underspecification of discourse conditions, which is directly connected to a parametric choice between the L1 and the L2 grammars (Sorace, 2011). If the speaker finds that there is no optional syntax in the L2, attrition is not expected to take place. Data from Tsimpli et al. (2004) suggest that near-native L2 English speakers of L1 Italian find anaphora resolution involving pronominal forms to be slightly problematic, showing residual optionality at advanced L2 stages, despite the L1 Italian preference for null subject pronouns. This same asymmetry was also confirmed in Sorace and Filiaci’s (2006) study, where overt subject pronouns tended to be overextended by near-native L2 Italian speakers of L1 English.

The general scope of the IH is over the syntax-pragmatics interface, particularly anaphora resolution in different languages. In order to check whether typological proximity had an effect on the difficulty associated with the use of pronominal subjects, Sorace, Serratrice, Filiaci, and Baldo (2009) tested two groups of bilingual children exposed to two different language scenarios: (a) Italian-English and (b) Italian-Spanish. These two language pairs were especially interesting because the authors could compare the interaction between a null subject language (NSL) and a non-null subject language (NNSL) in (a), and between two NSLs in (b). Their results indicated that both groups of bilinguals had a significantly higher acceptance of overt subjects referring to topic antecedents than monolingual Italian children. While this behavior had been predicted for the Italian-English bilingual group in light of the IH in previous studies, the overextension of overt pronouns in bilingual speakers of two NSLs (e.g. Spanish and Italian) could not simply be attributed to cross-linguistic influence (CLI), since both languages share a default null subject. Sorace (2011) considers the overextension of the overt pronoun to be a default strategy used to compensate for possible failure in mapping pronominal choice and pragmatic conditions. Under this account, the overextension of overt pronouns by Italian-Spanish bilinguals is caused either by obstacles in computing information in real time or by processing problems which are naturally generated when differentiating the two languages, i.e., due to a general effect of bilingualism. Recent research by Filiaci (2010) and Filiaci, Sorace, and Carreiras (2014) indicates that even though Spanish and Italian are both NSLs, they differ with respect to the scope of the overt pronoun, which suggests that CLI is also at play despite the typological similarity and apparently similar null subject distribution.
in the two languages. The differences between the Spanish and Italian pronominal systems (one vs. two series of 3rd person overt pronouns, respectively) may affect the properties of overt pronouns in these languages, especially in terms of the accessibility of their antecedents.

Regarding HL acquisition, Sorace (2005) claims that the input HSs receive is presumably attrited, qualitatively different or quantitatively impoverished, as their parents have likely undergone L1 attrition. Since the IH was originally proposed to cover issues raised in near-native acquisition, any unwarranted extension to other domains — such as HL acquisition — would lead to misinterpretations (Sorace, 2011). However, Montrul and Polinsky (2011) justify the application of the IH to HL acquisition in that the changes that take place in near-native acquisition should also take place in HL acquisition, precisely because of the quality of the input HSs receive. Thus, comparing HL to L2 acquisition can provide even richer support for the IH, despite the obvious differences between the two types of learners.¹

3. Null subject distribution in Portuguese

3.1 European Portuguese

With respect to the distribution of null subjects, EP has been claimed in the literature to be a consistent NSL of the Italian type (Rizzi, 1982; Jaeggli, 1984; Roberts & Holmberg, 2010). This means that EP, like Italian, has the following properties (Barbosa et al., 2005):

a) phonologically null subjects;

(4) a. Chegaram.
   ‘They arrived.’
   b. *Arrived.

b) SV, VS order alternations (“free inversion”)

(5) a. O João chegou.
   ‘John arrived’
   b. Chegou o João.

¹. A study by Keating, VanPatten, and Jegerski (2011), for instance, tested backward anaphora resolution in both HSs and L2 Spanish learners. While not directly testing the IH (as their target L2ers were highly proficient but not necessarily near-native) their results show that these HSs and L2ers display different behavior from one another and from monolingual Spanish speakers regarding the property in question, which in turn suggests no direct advantage for HSs despite early exposure to Spanish.
'John arrived'

c) lack of *that-trace effects

(6) a. Que candidato disseste que ganhou as eleições?
    ‘Which candidate did you say won the elections?’
b. *Which candidate did you say that won the elections?

In (4), the subject pronoun *Eles* “they” is replaced by a phonologically null subject, whose referent can be retrieved via contextual clues. Example (5) illustrates that both preverbal and postverbal subjects can surface in this language. In (6), we see that EP allows the subject to be extracted from after the complementizer *que*, ruling out *that*-trace effects.

Barbosa et al. (2005) also point out that NSLs, such as EP, have yet another property that distinguishes them from NNSLs. In these languages, there is generally no co-reference between overt embedded pronouns and matrix subjects. Costa, Faria, and Matos (1998) show that, especially in coordinated structures, an alternation between null and overt embedded subjects yields distinct co-reference patterns, as illustrated in (7):

(7) a. A Helena viu a Maria no cinema mas Ø não a cumprimentou.
    ‘Helenai saw Maria at the movie theater but did not greet her.’
b. A Helena viu a Maria no cinema mas ela não a cumprimentou.
    ‘Helenai saw Maria at the movie theater but she did not greet her.’

from Costa et al. (1998, p. 176)

The results discussed in Costa et al. (1998) suggest that, in EP, anaphora resolution strategies have a direct link with Chomsky’s Avoid Pronoun Principle, which states that “a lexical pronoun should be avoided whenever possible in favor of pro or PRO” (Chomsky, 1981, p. 65). In other words, EP speakers seem display the following behavior: if the coordinated subject is null, co-reference is automatically established at the highest c-commanding position in the clause, and in the case of an overt coordinated subject, the choice becomes less automatic given the different possible candidates for co-reference.

To sum up, it can be said that, with respect to the distribution of empty categories, EP does not differ significantly from consistent Romance NSLs such as Spanish and Italian. Regarding anaphoric contexts, when presented with a null subject, EP speakers choose to establish co-reference with a higher antecedent, but are faced with different possible referents to an overt pronominal subject, with a non-c-commanding antecedent being the best candidate.
3.2 Brazilian Portuguese

Unlike typical Romance languages, BP has undergone a series of changes with respect to subject pro-drop, especially in 3rd person contexts. Duarte (1993, 1995) suggests that BP is a language in transition from a NSL to a NNSL, due to the fact that the person-verb paradigm has been losing its uniformity over the last decades. Specifically, the 2nd person pronouns *tu* and *vós* (you, singular and plural, respectively) have been lost (or replaced with *você* and *vocês*) in many regions of Brazil. These new forms require agreement with the verb in the 3rd person, despite the fact that they are being used to address the hearer. Moreover, in most parts of Brazil where *tu* has remained, there is a mixed-agreement system where 2nd and 3rd person forms have been collapsed. This is taken as one of the reasons why in certain contexts an originally obligatory null subject pronoun has become optional. This can be seen in contexts with embedded subjects co-referential with the main subject of the matrix clause and left-dislocation of the subject as shown in (8a) and (8b):

(8) a. *Ela* ficou solteira porque *ela* quis.  
She stayed single because she wanted 
‘She remained single because she wanted to.’

b. A Clarinha {elai} cozinha que é uma maravilha.  
The Clarinha she cooks that is a wonder 
‘Clarinha she can cook wonderfully.’

With respect to Chomsky’s Avoid Pronoun Principle, BP certainly generates puzzling questions given the optionality shown for overt pronouns. In sentence (9), no focus reading or semantic/pragmatic effects are present. In addition, one finds optional alternation with null counterparts of the pronouns in bold.

(9) Porque *eu* / Ø não ‘tava certo se *eu* / Ø ia querer fazer escola técnica ou se *eu* / Ø queria continuar fazendo o científico.  
because I not was sure if I would want to make school technical or if I wanted to continue making the scientific 
‘Because I was not sure whether I wanted to go to technical school or if I wanted to continue high school.’

(9) Porque *eu* / Ø não ‘tava certo se *eu* / Ø ia querer fazer escola técnica ou se *eu* / Ø queria continuar fazendo o científico.  
because I not was sure if I would want to make school technical or if I wanted to continue making the scientific 
‘Because I was not sure whether I wanted to go to technical school or if I wanted to continue high school.’

(9) Porque *eu* / Ø não ‘tava certo se *eu* / Ø ia querer fazer escola técnica ou se *eu* / Ø queria continuar fazendo o científico.  
because I not was sure if I would want to make school technical or if I wanted to continue making the scientific 
‘Because I was not sure whether I wanted to go to technical school or if I wanted to continue high school.’

Kato (1999) sums up the BP system by stating that, despite the core grammar of modern BP having lost the null referential subject, it has retained the null expletive (10a), it allows for an arbitrary *pro* (10b) and it can have both a null bound pronoun (10c) and an “anaphoric” *pro* (10d). The examples below show these occurrences, respectively:
In sum, it has been claimed that BP differs from EP and from consistent Romance NSLs in the following respects: (a) weaker agreement in the person-verb paradigm yields more frequent occurrence of overt pronouns; (b) embedded subjects co-referential with the main subject of the matrix clause are not required to be dropped; (c) left-dislocation of the subject entails optionality of the overt pronoun; (d) when there is optionality, no focus reading or semantic/pragmatic effects are present.

4. Research questions and hypotheses

As previously mentioned, this study deals simultaneously with heritage acquisition of BP and adult L2 EP acquisition. The following research questions and hypotheses of the present study were formulated in light of the syntactic differences BP and EP and taking into account the research trends in heritage and adult L2 acquisition:

a. Do the HSs and L2 learners understand the syntactic differences between BP and EP and are they able to employ such knowledge in the present experiment in both BP and EP-modes?

We predict that dominance will yield an effect for HSs and L2 learners’ performance. This entails that heritage BP speakers, being dominant in EP and in light of the typical profile of HSs (see Benmamoun et al., 2013), should not pattern with BP controls in BP-mode, but should pattern with EP controls in EP-mode. Late L2 learners of EP, conversely, like native and BP-dominant speakers, should pattern with BP controls in BP-mode and may or may not differ from EP controls in EP-mode. It has been shown that successful acquisition of L2 syntax is likely to take place in naturalistic settings (see Isabelli, 2004; Rothman & Iverson, 2007;
Rothman, 2008), so it might be possible that BP speakers manage to acquire the EP patterns despite their late age of L2 onset;


We expect to find patterns of CLI caused by an influence of dominance, due to increased exposure to EP and decreased exposure to BP prior to ultimate attainment of the L1 grammar (Putnam & Sánchez, 2013). For instance, in the case of HSs, this will be seen if they pattern with EP controls in EP-mode and show evidence of EP influence in BP-mode. As for L2 learners, the prediction would be the exact opposite, with BP-like behavior in both modes. Alternatively, it is possible that there will be a general effect of bilingualism ruling out CLI, that is, where neither HSs nor L2 learners perform like the controls in either mode possibly due to an effect of a compounded state of mind affecting grammatical representations in bilinguals (e.g. Cook, 2003, 2006) or due to processing efficiency affected by bilingualism (see e.g., Kroll, Bobb, Misra, & Guo, 2008; Bialystok, Craik, Green, & Gollan, 2010). Given the pattern of overextension of overt pronouns by Italian-Spanish bilinguals found in Sorace et al. (2009), it is likely that this behavior will be found in the language scenario in this study, given the high typological proximity between BP and EP. In light of the IH, this behavior is triggered by obstacles in computing information in real time or by processing difficulties linked to a general effect of bilingualism (Sorace, 2011).

c. If there is CLI, is it conditioned by age of arrival in Portugal?

We predict that there will be an age effect for CLI. There are several scenarios that can be predicted to play out. First, one might expect that age delimits potential for acquisition of new syntax (DeKeyser, 2000; Bley-Vroman, 2009; Long, 2005, Granena & Long, 2013). If so, then the prediction would be that L2 learners should show evidence of BP influence in EP-mode as a factor of age. It should also be the case, if age is truly a factor, that HSs would not show evidence of CLI at all in either mode, since they are child L2ers of EP and native L1 speakers of BP. This, however, is complicated by what we know about the typical case of HS ultimate attainment and the hypothesis in (b) that dominance will modulate CLI. If, however, there is no critical period for the acquisition of syntax per se (e.g., Schwartz & Sprouse, 1994, 1996; White, 1989, 2003), and especially if in naturalistic settings L2ers have the best chance to acquire new syntax (e.g., Isabelli, 2004, Rothman & Iverson, 2007, Rothman, 2008), then L2 learners would be expected to show no CLI from BP to EP. They might, however, show some CLI from EP to BP if they have experienced L1 attrition and/or shift in dominance.
5. Methodological approach

5.1 Participants

The participants in this study were divided into four groups, labeled as follows:

- **L2 group—Late L2 learners**: the participants in this group (n=20) were born in Brazil and moved to Portugal in adulthood. At time of testing, their mean age was 37.9 years (range=25–58). They acquired EP as an L2 upon arrival in Portugal. The mean age of arrival for the group was 27.8 (range=19–43), and the mean length of time living in Portugal, which equates to length of exposure to EP, was roughly 10 years (range=6–30 years);

- **HS group—Heritage Speakers of BP**: these participants (n=17) have a mean age of 29.1 years (range=18–52) at the time of testing. All participants were born in Brazil but moved to Portugal at an early age (mean=5.6; range=1–8 years old). The mean length of time living in Portugal at time of testing was 23.5 years (range=14–45);

- **BPC—BP control group**: this group (n=20) includes native BP speakers tested in Brazil. Their mean age is 31 (range=18–54). The participants had little or no exposure to EP outside of access to multimedia shared across the countries;

- **EPC—EP control group**: this group (n=20) consists of native EP speakers with little or virtually no exposure to BP outside of access to multimedia shared across the countries. Their mean age at the time of testing was 30.5 years (range=20–52).

All participants recruited in Portugal (L2ers, HSs and EPCs) were tested in and around the city of Braga, in the Minho region of northern Portugal between March and July of 2014. The BPC group was recruited in the city of Fortaleza, in Northeast Brazil, during the months of July and August of 2014. All participants had some knowledge of a foreign language, mostly English and Spanish, but none claimed to have full proficiency in any language other than Portuguese. All the participants had normal or corrected vision and normal hearing, except for one participant who was achromatic.

All of the participants were asked to fill out a language background questionnaire in which they provided information about their level of formal instruction, with the goal of increasing homogeneity within each group. All participants had completed a High School degree, and most had some college education. The target groups, the L2ers and HSs, had a longer version of the questionnaire, where they indicated what other languages they were proficient in, before and after moving to Portugal. This was done in order to single out cases of different pronoun patterns due to possible transfer from other languages. Thus, only participants who
claimed to be proficient in EP and who had not reached advanced fluency in other languages acquired in naturalistic settings were considered. While no standardized tests were administered to measure their EP proficiency, the target groups were asked to indicate the frequency with which they use each variant, out of 100% (e.g., BP: 40%, EP: 60%). The average for each of the target groups was the following: L2ers–BP: 58%, EP: 42%; HSs–BP 20.59%, EP: 79.41%. This shows that all target participants were aware of the differences between the two languages despite their mutually intelligibility and typological similarity. These questionnaires also included questions about when they immigrated to Portugal, so as to divide them according to age of arrival, and whether they had lived in other parts of Brazil or Portugal, to account for possible dialectal influence.

5.2 Mode trigger

Bilinguals have been described in the literature to display different language modes in their everyday lives (see Grosjean, 1998, 2008, for discussion). This entails that the speaker switches from one language mode to another, depending on their interlocutor, which can have an impact on both production and perception. At times, bilinguals even resort to using language mixing such as code-switching and borrowing, when the interlocutor is also bilingual in the same languages (Grosjean, 1998). Since the mode the speaker is in translates to the state of activation of either language, the testing conditions were adapted to generate a mode trigger, which would indicate to the participants that they were going to be tested in BP or EP. This mode split was for us to check whether they respond differently to the task when in BP-mode versus EP-mode. Therefore, two versions of the same task were created that were only adjusted for vocabulary distinctions between the varieties. The testing sessions were counterbalanced so that half the participants were tested first in BP and then in EP (by native speakers of each variety), and the other half were tested in the reverse order. There was a minimum of one week between the two testing sessions, to avoid possible priming effects. We included a mode-trigger rapport at the beginning of both sessions, in order to ensure that the participants were aware of the difference between the BP setting and the EP setting. The native EP speaker who conducted the EP version of the tasks prepared this setting by asking them general questions related to their experience in Portugal, and what they liked the most about Portuguese culture, cuisine, etc. In BP-mode, they were asked to say a little about what they missed from their home country, and if/how they kept ties with Brazil, such as how often they visited, whether they participated in Brazilian events in the area and so on. They were considered ready for testing after a period of about five minutes.
5.3 Experiment

The experiment was designed to test the participants’ preferences for forward anaphora resolution. The bilingual groups performed both versions of the task (as described in 5.2) while the controls only completed the version corresponding to their L1. The experiment is a Picture Verification Task (PVT), adapted from Sorace and Filiaci (2006), where they tested anaphora resolution in Italian-English bilinguals. We have chosen this PVT because it targets anaphora resolution in potentially ambiguous contexts, which is where the PAH is shown to be less flexible.

The experimental items consisted of 15 sentences divided into three conditions: overt embedded subject (OES), null embedded subject (NES) and left-dislocated subject (LDS), as illustrated in (11), (12) and (13), respectively:

(11)  *A senhora acena para a garota enquanto ela atravessa a rua.*
    ‘The lady waves at the girl while she crosses the street.’

(12)  *O pai acena para o filho enquanto Ø anda de bicicleta.*
    ‘The father waves at the son while rides a bike.’

(13)  *O porteiro, ele cumprimenta o carteiro enquanto abre a porta.*
    ‘The doorman, he greets the mailman while opens the door.’

This task was built on a web platform called SurveyGizmo, which offers a paid subscription that allows the user to create and design experiments with pictures, audio and other media. Each test item consisted of a sentence shown on the computer screen with three pictures. Participants were asked to read the sentence out loud (to ensure full comprehension of all the words in the sentence) and then choose the picture that best described the sentence they had just read. The pictures showed three possible referents for the overt pronoun in the OES condition, for the null pronoun in the NES condition, and for the left-dislocated subject in the LDS condition. The LDS condition was included as a control condition only to ensure that all the participants interpreted that as a BP-only construction, as this construction is typically not found in native EP. Since there is also a null embedded subject in these items, the expected referent should be the matrix subject. For this reason, this condition was treated as a filler condition, not as an extension of the NES condition, and thus is not included in the analysis.

The three referents available in the target conditions were: (i) the matrix subject; (ii) the matrix object; (iii) a disjoint referent. All items were randomized, and within each item, the order in which the pictures appear was also be randomized to avoid priming effects. Participants had the possibility of choosing more than one referent in case of extreme doubt, so as not to force a choice when two options are equally possible. Once the participants made their choice, they clicked
on the button *Próximo* (next) to move on to the next item. All of their choices were automatically registered by the web platform after each click. Figure 1 illustrates a screenshot of one of the items from the PVT-BP version.

Figure 1. Screenshot of random item from BP version of the task

For both BP and EP, co-reference with the matrix subject is expected to be the preferred option in contexts with null embedded subjects, as is the case in other Romance NSLs such as Italian (Carminati, 2002; Sorace & Filiaci, 2006), Spanish (Filiaci, 2011; Bel & García-Alcaraz, 2015), and Catalan (Mayol & Clark, 2010), for instance. However, the differences between the null subject distributions in BP and EP give rise to potentially different co-reference preferences for the sentences with overt embedded subjects. In Figure 1, for instance, BPCs might accept all three readings, since the overt subject of the embedded clause can be co-referential with the matrix subject in this language. While these readings are also possible in EP, the PAH would predict that any reading other than co-reference with the object is highly marked, since overt pronouns prefer to have a referent lower in the phrase structure. EPCs were thus expected to choose the top left picture as the only possible reading for this sentence, confirming co-reference with the object.
6. Results

First, we offer a descriptive analysis of the performance of each group. The analysis of the results of this task was limited to the two main conditions, overt embedded subject (OES) and null embedded subject (NES). The left-dislocated subject (LDS) condition was used as a filler, and thus is not included in the statistical analysis. Figures 2 and 3 below show the distribution of each possible referent across all groups and in both target conditions.
Since external referents were virtually not selected, the data do not allow for any significant comparison. Thus, we have decided to include only subject and object referents in our analysis. Even though three referents were possible for all conditions, we believe that using the two most prominent co-reference options should suffice to clearly identify the different patterns shown by the groups in each mode and condition. Given the low frequency of items for which two referents were selected, we have chosen to code such items as two separate responses, as we are convinced that it does not affect the pattern of results. The overall percentage of acceptance in each of the target conditions across all groups when considering co-reference with the matrix subject as the baseline is presented in Figure 4. Figure 5 illustrates the overall choices taking co-reference with the object as the baseline.

![Figure 4](image1.png)  
**Figure 4.** Percent acceptance, co-reference with subject

![Figure 5](image2.png)  
**Figure 5.** Percent acceptance, co-reference with object
This analysis shows that the BP controls (BPCs) and EP controls (EPCs) seem to differ from one another in the OES condition. While the BPCs established co-reference with the subject in 38% of the cases, the EPCs set this preference at 13%. In BP-mode, the HSs preferred co-reference with the subject in 20% of the items, against 30% for the L2ers and 38% for the BPCs. When tested in EP-mode, the HSs chose subject co-reference 25.9% of the time, against 33% for the L2ers and 13% for the EPCs. Due to the low rates of co-reference with the subject, in this condition the preference seems to be for co-reference with the object. This is clear in both control groups (object co-reference: BPCs = 71%; EPCs = 94%). The L2ers in BP-mode chose co-reference with the object in 72% of the cases, versus 84% when in EP-mode. The HSs chose co-reference with the object at 77.6% in BP-mode, and 72.9% in EP-mode.

In the NES condition, the BPCs established co-reference with the subject in 81% of the cases, and EPCs, in 79% of the items. When tested in BP-mode, the HSs subject co-reference in 89.4% of the cases, versus 81% for both the L2ers and the BP controls. The EP-mode numbers indicate that the HSs preferred co-reference with the subject in 90.5% of the cases, compared to 77% for the L2ers and 79% for the EPCs. Co-reference with the object was also chosen in this condition, but at a lower rate (BPCs = 21%; EPCs = 27%). Across the two modes, the two target groups do not appear to show distinct preferences (L2ers: BP-mode= 27%, EP-mode=28%; HSs: BP-mode 10.5%, EP-mode: 9.4%).

**Statistical models**

In order to test whether the target groups displayed any statistical differences across the different modes and conditions, we used a mixed effects linear regression model. The threshold for statistical significance was set at $p \leq .05$. The statistical model included variables of group (BPC, L2ers, HSs, EPC), mode (BP vs. EP) and condition (OES vs. NES) as fixed effects.

If we consider the effect of overtness — establishing a comparison between the NES and OES conditions to see the effect of the inclusion of an overt subject — we reach the results illustrated in Tables 1 and 2.

Considering co-reference with the subject as the baseline, our analysis shows that:

a. BPCs are statistically different from EPCs ($p=0.001$);
b. The L2 learners in EP-mode are also different from EPCs \( (p=0.002) \), but not different from BPCs when in BP-mode \( (p=0.509) \), nor are they different from themselves across the modes \( (p=0.699) \); 

c. The HSs in EP-mode pattern with the EPCs \( (p=0.822) \), but in BP-mode they differ from the BPCs \( (p=0.002) \). They do not differ from themselves across the modes \( (p=0.716) \), so no mode effect was found for either L2 or HS groups; 

d. The HSs are significantly different from L2ers in EP-mode \( (p=0.003) \), and in BP-mode \( (p=0.014) \).

Table 1. Minimal adequate model of mixed effects linear regression of anaphora resolution preferences using co-reference with the subject as the baseline

<table>
<thead>
<tr>
<th>Random effects</th>
<th>Group</th>
<th>Name</th>
<th>Variance</th>
<th>SD</th>
<th>Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>informant</td>
<td>(Intercept)</td>
<td>0.8711</td>
<td>0.9333</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP-mode</td>
<td>2.9468</td>
<td>1.7166</td>
<td>-0.64</td>
</tr>
<tr>
<td></td>
<td>stimulus</td>
<td>(Intercept)</td>
<td>0.5688</td>
<td>0.7542</td>
<td></td>
</tr>
</tbody>
</table>

Observations: 1140; Groups: Informant: 77, Stimulus: 10

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Variable</th>
<th>Estimate</th>
<th>SE</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Intercept)</td>
<td>-26257</td>
<td>0.5871</td>
<td>-4.472</td>
<td>7.74e-06</td>
</tr>
<tr>
<td></td>
<td>HS/BP-mode</td>
<td>0.8842</td>
<td>0.6124</td>
<td>1.444</td>
<td>0.148756</td>
</tr>
<tr>
<td></td>
<td>HS/EP-mode</td>
<td>1.1037</td>
<td>0.6477</td>
<td>1.704</td>
<td>0.088390</td>
</tr>
<tr>
<td></td>
<td>L2ers/BP-mode</td>
<td>1.5588</td>
<td>0.5780</td>
<td>2.697</td>
<td>0.006999</td>
</tr>
<tr>
<td></td>
<td>L2ers/EP-mode</td>
<td>1.6328</td>
<td>0.6130</td>
<td>2.664</td>
<td>0.007727</td>
</tr>
<tr>
<td></td>
<td>BPC</td>
<td>2.0039</td>
<td>0.5732</td>
<td>3.496</td>
<td>0.000472</td>
</tr>
<tr>
<td></td>
<td>Condition: NES</td>
<td>4.5011</td>
<td>0.7155</td>
<td>6.291</td>
<td>3.15e-10</td>
</tr>
<tr>
<td></td>
<td>HS/BP-mode:NES</td>
<td>-0.1203</td>
<td>0.7366</td>
<td>-0.163</td>
<td>0.870312</td>
</tr>
<tr>
<td></td>
<td>HS/EP-mode:NES</td>
<td>0.1668</td>
<td>0.7449</td>
<td>0.224</td>
<td>0.822832</td>
</tr>
<tr>
<td></td>
<td>L2ers/BP-mode:NES</td>
<td>-1.6721</td>
<td>0.6526</td>
<td>-2.562</td>
<td>0.010397</td>
</tr>
<tr>
<td></td>
<td>L2ers/EP-mode:NES</td>
<td>-1.8843</td>
<td>0.6130</td>
<td>-2.977</td>
<td>0.002913</td>
</tr>
<tr>
<td></td>
<td>BPC:NES</td>
<td>-2.0246</td>
<td>0.6499</td>
<td>-3.115</td>
<td>0.001838</td>
</tr>
</tbody>
</table>

Reference levels of categorical fixed effects: Group: EPC; Condition: OES
Log likelihood: -526.1; Deviance: 1052.2; Residual degrees of freedom: 1124

2. The numbers shown here take EPC as the intercept for comparison. All values not depicted in Tables 1 or 2, namely those comparing the other three groups, could only be obtained by changing the intercept.
When using co-reference with the object as the baseline for the analysis, the results still hold, as these two co-reference options were, for the most part, complementary:

a. BPCs are statistically different from EPCs (p=0.003);
b. L2 learners in EP-mode differ from EPCs (p=0.030), but not from BPCs when in BP-mode (p=0.493). Across the two modes, they also do not show any statistical difference (p=0.099);
c. HSs in EP-mode behave like the EPCs (p=0.496), but differ from the BPCs when in BP-mode (p=0.039). They also show similar behavior across the two modes (p=0.647);
d. L2ers and HSs are significantly different from one another both in EP-mode (p=0.013), and in BP-mode (p=0.007).

Table 2. Minimal adequate model of mixed effects linear regression of anaphora resolution preferences, using co-reference with the object as the baseline

<table>
<thead>
<tr>
<th>Random effects</th>
<th>Group</th>
<th>Name</th>
<th>Variance</th>
<th>SD</th>
<th>Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>informant</td>
<td>(Intercept)</td>
<td>0.9057</td>
<td>0.9517</td>
<td></td>
</tr>
<tr>
<td></td>
<td>stimulus</td>
<td>(Intercept)</td>
<td>0.3977</td>
<td>0.6306</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP-mode</td>
<td>3.8833</td>
<td>1.9706</td>
<td>−0.77</td>
</tr>
</tbody>
</table>

Observations: 1140; Groups: Informant: 77, Stimulus: 10

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Variable</th>
<th>Estimate</th>
<th>SE</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Intercept)</td>
<td>3.5952</td>
<td>0.6575</td>
<td>5.468</td>
<td>4.55e−08</td>
</tr>
<tr>
<td></td>
<td>HS/BP-mode</td>
<td>−2.0775</td>
<td>0.6990</td>
<td>−2.972</td>
<td>0.002958</td>
</tr>
<tr>
<td></td>
<td>HS/EP-mode</td>
<td>−2.2114</td>
<td>0.7348</td>
<td>−3.009</td>
<td>0.002618</td>
</tr>
<tr>
<td></td>
<td>L2ers/BP-mode</td>
<td>−2.4324</td>
<td>0.6770</td>
<td>−3.593</td>
<td>0.000327</td>
</tr>
<tr>
<td></td>
<td>L2ers/EP-mode</td>
<td>−1.4711</td>
<td>0.7254</td>
<td>−2.028</td>
<td>0.042555</td>
</tr>
<tr>
<td></td>
<td>BPC</td>
<td>−2.4641</td>
<td>0.6779</td>
<td>−3.635</td>
<td>0.000278</td>
</tr>
<tr>
<td></td>
<td>Condition: NES</td>
<td>−5.0051</td>
<td>0.7434</td>
<td>−6.733</td>
<td>1.66e−11</td>
</tr>
<tr>
<td></td>
<td>HS/BP-mode:NES</td>
<td>0.8929</td>
<td>0.7985</td>
<td>1.118</td>
<td>0.263440</td>
</tr>
<tr>
<td></td>
<td>HS/EP-mode:NES</td>
<td>0.5440</td>
<td>0.8006</td>
<td>0.679</td>
<td>0.496845</td>
</tr>
<tr>
<td></td>
<td>L2ers/BP-mode:NES</td>
<td>2.5374</td>
<td>0.7243</td>
<td>3.503</td>
<td>0.000459</td>
</tr>
<tr>
<td></td>
<td>L2ers/EP-mode:NES</td>
<td>1.5813</td>
<td>0.7296</td>
<td>2.167</td>
<td>0.030207</td>
</tr>
<tr>
<td></td>
<td>BPC:NES</td>
<td>2.1756</td>
<td>0.7366</td>
<td>2.953</td>
<td>0.003143</td>
</tr>
</tbody>
</table>

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 1
Reference levels of categorical fixed effects: Group: EPC; Condition: OES
Log likelihood: −516.5; Deviance: 1032.9; Residual degrees of freedom: 1124
7. Discussion

This task took into consideration the aforementioned differences between BP and EP to test the extent to which: (a) HSs and L2 learners understand and employ their knowledge of the syntactic differences between BP and EP in the present experiment in both BP and EP-modes, (b) there is evidence of cross-linguistic influence (CLI) in any direction: BP⇒EP, EP⇒BP and/or BP⇔EP, and (c) possible CLI is conditioned by age of arrival in Portugal. As shown in Section 3, BP and EP are reportedly different regarding the distribution of null subjects. In the event of an overt embedded subject, both languages display three possible co-reference options, namely with the subject of the matrix clause, with the object of the matrix clause, or some 3rd person referent outside the clause (i.e. disjoint reference). The difference between BP and EP in this case is that since EP speakers generally prefer to have a null embedded subject in contexts where co-reference with the matrix subject is intended, the presence of an overt embedded subject should indicate pragmatic effects, such as focus or contrast. As shown in Duarte (1993, 1995), overt embedded subjects in BP do not carry these pragmatic effects, and thus can freely appear in place of a null pronoun while the referent remains the same.

It must be noted that all co-reference options are grammatical and the target anaphora resolution is highly dependent on preference. In contexts with a null embedded subject, as predicted by the PAH, co-reference is established at a higher position in the clause (subject of the matrix clause), and this was confirmed for both BPC and EPC control groups. In contexts with an overt embedded subject, both languages are predicted to set co-reference lower in the clause, namely with the object of the matrix clause. This is also confirmed for both BPC and EPC control groups. Nevertheless, there are some differences apparent in the control data presented. The statistical difference we see between BPCs and EPCs in the OES condition indicates that, even though BPCs, like EPCs, prefer co-reference with the object, they still allow for co-reference with the subject significantly more than EPCs (BPC=38%; EPC=13%). This difference is expected given the optionality of the overt pronoun in BP (Duarte 1993, 1995). While EP speakers make use of overt pronouns for pragmatic reasons such as emphasis or contrasts, BP speakers are argued to show optionality of overt pronouns without the same effects. Thus, the presence of the overt pronoun does not change co-reference in BP to the same extent that it does in EP. We now proceed with the discussion of the results taking into consideration each research question presented in Section 4.

a. do the HS and L2 learners understand the syntactic differences between BP and EP and are they able to employ such knowledge in the present experiment in both BP and EP-modes?
With respect to HL and L2 acquisition, we made the prediction that dominance would play a role in the performance of the target groups in this experiment. The L2ers were expected to pattern with the BP controls in BP-mode, as dominant native BP speakers. This was indeed the case. Despite what previous research has found about successful L2 acquisition of syntax in naturalistic settings (Isabelli, 2004; Rothman, 2008), this did not seem to be the case for BP natives exposed to EP in Portugal, because anaphora resolution does not involve purely syntactic computations, but a combination of features that lie at the syntax-pragmatics interface. Since we also see that these speakers do not display EP-like behavior in EP-mode, we conclude that they do not have the same preferences as native EP speakers, maintaining their BP-like behavior. This is in line with the IH, as constructions involving the syntax-pragmatics interface appear to be problematic in L2 acquisition (Sorace & Filiaci, 2006; Sorace, 2011).

This influence from the native language on the second is known as L1 transfer (see e.g., Ellis, 2006; Schwartz & Sprouse, 1994, 1996, among many others). Since L2ers do not have to acquire a new syntactic structure, their preferences are much less challenged. In the case of Spanish speakers learning Italian, for instance, they would be exposed to a new grammar, which involves learning a new set of rules. It has been shown that negative transfer is likely to occur to a much greater extent between closely related languages than between languages that are typologically distinct (Wolfram & Schilling-Estes, 1998; Long, 2007). In the case of two typologically similar languages like Spanish and Italian, learners acquire a new set of rules and typically transfer from the L1, but the similarities between the two languages could potentially delay the acquisition of these morpho-syntactic distinctions. It is therefore possible that the high degree of typological proximity and mutual intelligibility between BP and EP leads these learners to maintain their BP preferences, as they may not feel the need to learn a new set of rules.


Since the differences between the two languages lie at the syntax/discourse interface, the IH predicts that L2ers could potentially display some signs of attrition when in BP-mode (Sorace, 2000; Tsimpli et al., 2004), but this did not turn out to be the case. Naturally, attrition can only take place if the speakers fully acquire the features of the L2 that in turn affect their L1 (Sorace, 2011), and since there is no successful acquisition (as discussed with the previous research question), there is no possibility of attrition. Tsimpli et al. (2004) showed that anaphora resolution was slightly problematic for near-native L2 English speakers of L1 Italian, showing residual optionality at advanced L2 stages, even though L1 Italian monolinguals had a strong preference for null pronouns. In their study, attrition was found
for the interpretation of overt pronominal subjects among L1 speakers who were near-native L2 English speakers. We do not find the same pattern in these BP-EP bidialectal bilinguals precisely because there is no acquisition of new syntactic structures, and the difference between the two languages, as confirmed by the control data, lies in pragmatic preferences. In other words, co-reference of an overt embedded subject with the matrix subject in EP is not ruled out, but typically linked to pragmatic cues such as focus or contrast. In BP, the pattern seen indicates that these pragmatic effects are indeed not present, as they allow for co-reference with the subject in this context significantly more than the EPC, even though they still prefer lower co-reference as predicted by the PAH. Tsimpli et al. (2004) also showed that the directionality of cross-linguistic effects is linked to underspecification of interpretable features — the less restrictive language influences the other, but not vice-versa. In their study, English affects Italian, regardless whether it is the L1 or the L2, because it has the most economical syntax-pragmatics interface system for subject pronouns as it lacks null subjects. In Italian, much like in EP, the interface is more complex, since null subjects co-exist with overt subjects depending on pragmatic constraints. If the grammar of BP has reached a point where it is no longer a NSL, but rather a partial-NSL (Holmberg, Nayudu, & Sheehan, 2009) with a higher occurrence of overt subjects, the directionality of CLI should also be from BP to EP, and not the other way around. We see that L2ers in the NES condition maintained their BP preferences, which confirms the expected directionality of CLI.

The HSs displayed EP-like behavior in both modes in the OES condition. This can be linked to the dominance factor, as these speakers had been raised and educated in Portugal, using BP only in family contexts. It has been shown that dominance may have a stronger effect than age of acquisition with respect to language transfer. Montrul and Ionin (2012) argue that transfer from the stronger language is not only likely to occur, but also helps shape both L2 and adult heritage grammars. In the present study, no tests were carried out to measure the participants’ dominance, but during our initial rapport, these speakers claimed to pass as native EP speakers, though some admitted to an occasional use of Brazilian vocabulary in full EP conversations. We understand by their performance in this task that dominance seems to have an effect on these speakers. Even though the social context in Portugal is such that these speakers continue to have exposure to their L1 through mass media, they still live in an environment where they often interact with EP speakers and use EP to carry out their daily routines. Therefore, after an average of 23.5 years of EP exposure, it is natural that they are EP-dominant and thus display EP-like preferences for anaphoric reference in the OES condition. In the NES condition, HSs did not differ from either control group, accepting co-reference with the matrix subject as expected and in line with the PAH. The difference between
Comparing anaphora resolution

the HSs and the EPCs in the NES condition is not reported as significant, but it does show that these heritage speakers respected the PAH slightly more than the EPCs, thus establishing co-reference with the matrix subject in this condition more often (HS=90.5%; EPC=79%). Since these participants displayed EP-like behavior in both modes, this could be seen as evidence of CLI. The directionality of CLI is from the L2 to the L1 — from EP to BP, their heritage language — and this can be seen especially in contexts with overt embedded subjects. This indicates a difference between L2ers and HSs regarding the directionality of CLI, which we believe to be a consequence of EP dominance.

Studies supporting the IH have shown that bilinguals who speak a NSL and a NNSL have a tendency to overgeneralize overt pronouns, even in contexts where a null pronoun is preferred or required in the monolingual grammar (Tsimpli et al., 2004; Sorace & Filiaci, 2006). Keating et al. (2011) conclude that HSs do not perform better than L2ers regarding backward anaphora resolution in Spanish, but that both groups differ from monolingual counterparts. Studies investigating anaphora resolution in Italian have found that it can be affected whether the other language is a NSL or a NNSL, since both English-Italian and Spanish-Italian bilinguals performed differently from monolingual Italian speakers (Sorace et al., 2009). In this case, the IH predicts that bilinguals are not as efficient as monolinguals with respect to processing, and the differences found between bilinguals and monolinguals is a result of bilingualism itself, since these differences do not stem from the structure of the other language (Sorace, 2011). Thus, in light of the IH, the pattern shown by the HSs in the OES condition would not be a clear indicator of CLI, but possibly a result of processing difficulties, or a byproduct of their disambiguation strategy. However, in their study on Spanish-Italian bilinguals, Filiaci et al. (2014) argue that CLI should not be entirely ruled out as the overextension of overt pronouns in Italian by Spanish speakers can be justified by a combination of both CLI and the use of a default processing strategy. In the present study, HSs had in fact had significant BP input before being exposed to EP, as they were, on average, almost 6 years old when they arrived in Portugal. Moreover, Costa and Ambulate (2010) showed that EP monolingual children find anaphora resolution problematic in contexts with overt embedded subjects until age 5, preferring co-reference with the matrix subject, which is not expected in the adult grammar. This means that, even though the presence of the L1 can yield delays and persisting optionality in the acquisition of the L2, other factors may also play a role. If EP speaking children start setting their anaphora resolution preferences after age 5, and these HSs are going from a BP-only environment to a bidialectal BP-EP one precisely after age 5 (mean AoA=5.6), we can see that this is a crucial period for them as well. HSs first acquire BP anaphora resolution preferences and then reset
their preferences at the same pace as the EP monolingual children, resulting in EP-like performance in adulthood.

With respect to the differences between bilinguals and monolinguals, Sorace (2011) suggests that syntactic processing is less automatic for bilinguals than it is for monolinguals. This is seen as one of the reasons for bilinguals’ less effective processing of structures at the syntax-pragmatic interface, presumably due to limited access and integration of syntactic knowledge. This reduced ability to integrate syntax and contextual clues is reported in studies targeting processing strategies (see Kilborn, 1992; Roberts, Gullberg, & Indefrey, 2008). The overall conclusion of these studies is that bilinguals display a much larger effort in combining syntactic information with discourse conditions in real-time language use. Sorace (2011) claims that the overextension of overt pronouns in the bilingual grammar is a default strategy to compensate for this processing deficiency. In other words, overt pronouns are seen as a mechanism to reduce possible ambiguity. This behavior is also confirmed in our data, as HSs do set more flexible co-reference in the OES condition than the monolingual EP controls, allowing the overt pronoun to have the matrix subject as a co-referent more often. Even though this difference is not significant, it indicates a trend that could be confirmed by a larger sample size.

c. If there is CLI, is it conditioned by age of arrival to Portugal?

In the present study, our data show that age does in fact determine directionality of influence between BP and EP. The prediction was that L2ers would show BP effects in EP-mode, and that the reverse behavior would be found among HSs due to early exposure to both varieties. The L2ers do show a pattern that is more EP-like in both the OES and NES conditions when tested in EP-mode, but still behave significantly differently from monolingual EP speakers in the OES condition. The behavior displayed by the HS group seems to be a clear indicator of CLI, as signs of their EP dominance can be seen when they are tested in BP-mode, more specifically in the context of overt pronouns.

We see that the high degree of typological proximity between BP and EP seems to be the reason for these novel results. Even though both L2ers and HSs claimed to understand the two language modes, the results indicate that they show only one system, with language dominance as the factor that determines how close this system is to that of a monolingual speaker. Their anaphora preferences indicate that the BP-EP situation in Portugal is still one where the two languages coexist, but speakers make use of only one of them, despite the grammatical differences argued in the literature.
8. Conclusion

In this paper, we have analyzed comprehension data from an anaphora resolution task in Brazilian Portuguese (BP) and European Portuguese (EP). The results of this study suggest that the high level of typological proximity and mutual intelligibility between BP and EP hinders the acquisition of the EP anaphora co-reference preferences by Brazilian adult immigrants in Portugal. Instead, what we see in these late L2 learners, reflected both in BP and EP-mode, is the typical BP-like pattern, where anaphora co-reference is more flexible, presumably due to the optionality of overt subject pronouns in this language and the lack of the pragmatic effects that are present in EP. While Tsimpli et al. (2004) found attrition effects in L1 Italian-L2 English bilinguals, the same behavior is not found in this study, precisely because there must be acquisition of new grammar rules for L1 attrition to take place, and that was not the case for the L2 group.

Heritage speakers of BP in Portugal, on the other hand, seem to fully acquire the EP anaphora resolution preferences, and display EP-like behavior in both modes in the context of an overt embedded subject. The IH predicts that bilinguals have a tendency to overgeneralize overt pronouns, as seen with English-Italian (Sorace et al., 2009) and Spanish-Italian (Filiaci, 2010). This behavior was also seen in our HS data, though not to a significant extent. Instead, they seem to have acquired the BP distribution early on, and later shifted their preferences, as Costa and Ambulate (2010) show to be the case for monolingual EP children. We believe that this shift is due to increased exposure to EP, and consequently, decreased exposure to BP, and it appears that language dominance plays a significant role in the shaping of the new grammar, as discussed in Montrul and Ionin (2012).

Even though the syntax of the two languages is argued to be substantially different with respect to the distribution of null subjects, anaphora resolution involves other factors that lead these Brazilians to treat them as one, be it BP for the late learners, or EP for the heritage speakers. We intend to further pursue whether other domains of the grammar where the two languages display strong distinctions, such as the distribution of phonetically null objects, are more vulnerable. This will help us see whether a mode split can take place in some domains but not others, or whether BP and EP are treated as the same language across the board.

Acknowledgements

We would like to thank our participants for taking part in the studies and providing insightful ideas for future research. Many thanks to Telma Oliveira, Cristina Flores and Pilar Barbosa for helping facilitate data collection in Portugal. We are also grateful to Martin Corley and Björn
Lundquist for help with statistical analysis of the data. This study was partly funded by the HSL faculty at UiT — The Arctic University of Norway.

References


Comparing anaphora resolution


Comparing anaphora resolution


Slabakova, R. (2013). Adult second language acquisition: A selective overview with a focus on the learner linguistic system. Linguistic Approaches to Bilingualism, 3(1), 48–72. doi:10.1075/lab.3.1.03sla


El presente estudio examina la resolución de la anáfora en dos grupos de hablantes que han estado en contacto con el portugués brasileño y el europeo (BP y EP, respectivamente), teniendo en cuenta las diferencias en lo que se refiere a la distribución del sujeto nulo en estas dos variantes del portugués. La pregunta que nos proponemos contestar es si los bilingües tardíos BP-EP (edad de inicio: 29,1) y los hablantes de herencia de BP que crecen en Portugal (edad de inicio 5,6), se comportan de la misma forma que los nativos del grupo de control cuando realizan pruebas en ambos dialectos o si, por el contrario, muestran efectos del EP en su BP nativo o viceversa. Para la Interface Hypothesis (IH) esta es una pregunta interesante dado que esta hipótesis afirma que las interfaces externas deben estar sujetas a efectos generales de bilingüismo independientemente de la combinación de lenguas y edad (Sorace, 2011). Los resultados muestran que la edad tiene un efecto, puesto que los hablantes de herencia no se comportan como los bilingües tardíos, y muestran también que el alto grado de proximidad tipológica entre las dos lenguas descarta la posibilidad de adquisición bidialectal tardía.

Palabras clave: sujetos nulos, portugués, bilingüismo, bidialectismo, atrición

Authors’ addresses

Tammer Castro
CASTL / Department of Language and Linguistics
Faculty of Humanities, Social Sciences and Education
UiT — The Arctic University of Norway
TEO-H4 4.359
N–9037 Tromsø, Norway
tammer.castro@uit.no

Jason Rothman
School of Psychology & Clinical Language Sciences
University of Reading
Harry Pitt Building
Earley Gate
Reading, UK
j.rothman@reading.ac.uk
Marit Westergaard  
CASTL / Department of Language and Linguistics  
Faculty of Humanities, Social Sciences and Education  
UiT — The Arctic University of Norway  
N–9037 Tromsø, Norway  
marit.westergaard@uit.no

About the authors

Tammer Castro is a PhD candidate at UiT, The Arctic University of Norway. He is an affiliated member of the Tromsø branch of the project Bilingualism Matters, as well as a member of the research groups LAVA (Language Acquisition, Variation and Attrition), and AcqVa (Acquisition, Variation and Attrition). His research interests are first and second language acquisition, language attrition, bidialectalism and Portuguese morphosyntax.

Jason Rothman is a Professor of Multilingualism and Clinical Language Sciences at the University of Reading as well as a Professor II of Linguistics at UiT, The Arctic University of Norway. He has conducted research on first, second and third language acquisition as well as heritage language bilingualism, with a particular focus on the acquisition of morphosyntax. His research makes use of a diverse range of behavioral, psycho- and neurolinguistic methodologies to investigate quantitative and qualitative aspects of linguistic competence in different groups of learners.

Marit Westergaard is a professor of English Linguistics at UiT, The Arctic University of Norway and she also holds an adjunct professorship at NTNU Norwegian University of Science and Technology. Her research interests span monolingual and bilingual first language acquisition, second/third language acquisition, language variation and diachronic change, and more recently also heritage languages and language attrition.