

Comparing the Outcomes of Early and Late Acquisition of European Portuguese: An Analysis of Morpho-syntactic and Phonetic Performance

Cristina Flores, Universidade do Minho, Braga
Esther Rinke, Goethe Universität, Frankfurt
Anabela Rato, University of Toronto

ABSTRACT

The present paper compares the linguistic competence of German-Portuguese bilinguals with upper-intermediate German L2 learners (L2ers) of EP (European Portuguese) and with monolingual Portuguese speakers. The bilingual speakers are heritage speakers (HSs), who were raised bilingually with EP as the minority language and German as the majority language. The aim of our comparison is to verify in which way different input sources and maturational effects shape the speakers' linguistic knowledge. The findings of two studies, one focused on the morpho-syntactic knowledge of clitics and the other on global accent, corroborate the assumption that L2ers and HSs behave differently, despite superficial similarities observed in the morpho-syntactic study. In contrast to that of the L2ers', the accent of the HSs is perceived as being native-like, whereas their morpho-syntactic competence is mainly shaped by their dominant exposure to colloquial Portuguese and reduced contact with formal registers.

Keywords: *Heritage speakers, L2 learners, acquisition of clitics, global accent rating, European Portuguese, German*

1. INTRODUCTION

According to a large body of studies on simultaneous bilingual language acquisition (see, e.g., the overview in De Houwer, 1995), children who are exposed to two or more languages from birth acquire these languages qualitatively in a similar way to monolingual children. Nevertheless, many studies on the linguistic competence of second-generation immigrant adults, so called "heritage speakers" (HSs) who have acquired a minority language in the context of a dominant environmental language, claim that they typically show competence differences in their home language when compared to age-matched monolinguals, although they have learned this language from an early age (see Montrul, 2016, for a recent discussion). That *divergent* knowledge must not inevitably be equaled to *deficient* or *incomplete* knowledge is a growing demand within the research field of bilingual language development (e.g., Kupisch & Rothman, 2016; Pascual y Cabo & Rothman, 2012; Putnam & Sánchez, 2013). Various factors may contribute to variable competence outcomes in HSs, namely different input conditions, different underlying learning mechanisms, language dominance and several social factors such as cultural differences, ethnicity or birth order (see discussion in Hoff, 2006).

A way of determining the effects of influential factors on bilingual language development involves comparing early bilinguals with bilingual populations that differ with respect to some variables. This is the case of late L2ers who diverge from HSs in terms of age of onset of language acquisition (AoA) and type and quantity of language exposure. Several studies on HSs have compared these speakers with late L2ers in different linguistic domains (Au, Knightly, Jun, & Oh, 2002; Au, Knightly, Ju, Oh, & Romo, 2008; Cuza, & Frank, 2015; Keating, VanPatten, & Jegersky, 2011; Knightly, Jun, Oh, & Au, 2003;

Kupisch, 2012; Montrul, 2010, 2011; Montrul, Foote, & Perpiñán, 2008; O'Grady, Lee, & Choo, 2001; Santos & Flores, 2016).

Montrul (2010), for instance, claims that heritage bilinguals acquire the language from early on like monolingual speakers and therefore generally control phonological and syntactic features of their heritage language that are acquired at an early age. They usually acquire their knowledge through oral and naturalistic input. In contrast, L2ers typically learn their second language later in life, have access to a varying amount of written and aural input in the classroom, and are experienced with formal instruction. Nevertheless, according to the author, both L2ers and HSs may show developmental and transfer errors and variable proficiency. The differences between the two groups can be traced back both to different input conditions and different AoA. The commonalities result from the influence of the dominant language and the reduced input in comparison to monolingual speakers.

The aim of the present study is to understand the similarities and differences between L2ers and heritage bilinguals and to detect the effects of maturational and input factors in the performance of the two groups of German and EP speakers. The effects of differences on AoA and input between HSs and L2ers can be assessed in two ways: by examining properties that are shaped very early and represent a challenge for late learners (e.g., pronunciation), and by comparing properties that are stabilized through varied input sources (e.g., vernacular versus formal language registers). We report the results of two independent studies with heritage bilinguals, L2 learners and monolingual speakers of EP: i) an untimed written acceptability judgment task (AJT) concerning the use of clitics in EP and ii) a global accent rating task in which speech samples of the three speaker groups were rated for degree of nativeness by native Portuguese speakers. Particularly, the objectives of the research were to investigate i) if early and naturalistic exposure of the HSs to aural input as opposed to the late and more formal input of L2 speakers is reflected in their morpho-syntactic competence and ii) whether HSs early exposure to naturalistic input leads to a native-like pronunciation and is thus perceived as being distinct from that of late L2ers.

2. EXISTING PROPOSALS AND HYPOTHESES

Despite the long debate on the validity of the effects of maturational factors on language development, it is difficult to refute that differences between early and late language acquisition are related to AoA (c.f., Bley-Vroman, 1990; DeKeyser, 2000; Hyltenstam & Abrahamsson, 2003; Long, 1990; Meisel, 2011, among many others). These authors propose that adult language learning and child language acquisition are different because maturational factors constrain the process of language acquisition. Within this age-centered perspective, UG-based theories claim that children rely mainly on Universal Grammar (UG) and domain-specific learning processes, whereas adults learn a second language mainly on the basis of their L1 and general problem-solving strategies (Bley-Vroman, 1990). However, both for children and adults, the universal properties of UG also play a role in second language acquisition. To what extent this is the case is much debated in L2 research (in terms of full/partial/no access to UG and full/partial/no transfer of L1) but that discussion is beyond the scope of this paper (see Meisel, 2011, for a discussion).

If AoA is indeed crucial to shaping various types of language acquisition, it operates not only with regard to monolingual first language acquisition (L1) versus L2 acquisition but also to early bilingual language acquisition (2L1) versus late second language acquisition (L2). The application of this principle to all groups does not necessarily imply that early bilinguals, or specifically HSs, always perform like monolingual speakers in all respects. Some studies have shown that, in some domains of linguistic knowledge, heritage bilinguals are comparable to L2ers (Montrul, 2010, 2011; Montrul, Davidson, de la Fuente, & Foote, 2014). They may even be outperformed in some tasks by advanced L2ers that have acquired the language as adults. This possibility is typically shown in studies that compare heritage and L2 speakers in different experimental situations (e.g., Alarcón, 2011; Bowles, 2011; Montrul, 2011; Montrul, Foote, & Perpiñan, 2008; Montrul et al., 2014). Montrul, Foote and Perpiñan (2008), for instance, with respect to the acquisition of gender in Spanish, demonstrate that heritage bilinguals perform in a more target-like manner in oral production, whereas L2 learners perform better in written tasks. The authors argue that “accuracy scores on the oral task ... [are] more representative of fast, implicit, and automatically processed knowledge (typically acquired early in childhood), whereas accuracy scores of the written tasks could be taken to reflect ability with metalinguistic, explicit knowledge (typically acquired later)” (p. 541). Independently of various learning mechanisms (Bowles, 2011), the differing performance observed in oral and written tasks can also be attributed to the input available to the speakers: heritage bilinguals are mainly exposed to the spoken vernacular, whereas L2 learners usually have contact with the standard variety of the language and are more experienced with the written register. This observation is not surprising given that HSs are minority language speakers, who, in general, are educated in the majority language and have reduced access to formal instruction in their heritage language (HL). Pires and Rothman (2009), for instance, argue that “certain cases that could initially appear to be the outcome of incomplete acquisition in fact involve previous diachronic changes to the dialects that serve as the primary linguistic data of HSs” (p. 230).

Based on these remarks, we advance our first hypothesis.

Hypothesis 1: Heritage bilinguals will be less proficient than L2 learners regarding properties of the formal register that are infrequent in aural input or need written support or instruction in order to be stabilized. Regarding properties that are primarily acquired on the basis of naturalistic aural input, HSs perform in a more native-like manner than L2ers.

Extending this hypothesis to the comparison of linguistic modules, heritage bilinguals would perform better in linguistic domains that are shaped by aural input, such as phonological and phonetic competence, as already suggested by several researchers (e.g. Au et al., 2002; Chang, Yao, Haynes, & Rhodes (2011); Knightly et al., 2003; Oh, Jun, Knightly, & Au, 2003). However, this does not imply that HSs always achieve native-like pronunciation in their HL (see, for instance, Kupisch et al., 2014). Despite their overall ability to perceive and produce phonetic contrasts of their HL, HSs may produce accented speech due to cross-linguistic influence from the dominant language such as contrasts that are similar (but not identical) in both their languages. This is shown, for example, in Godson’s (2004) findings on the pronunciation of Armenian vowels by Western Armenian HSs living in the U.S., since only the Armenian vowels closest to English (the HSs’ dominant language) showed cross-linguistic influence. Significantly,

this influence did not fully neutralize the analyzed contrasts but rather maintained them apart. In addition, Chang et al. (2011), who analyzed HSs of Mandarin living in an English-dominant environment, report their near-native-like ability to produce various phonetic and phonological contrasts of their HL.

Further research on the comparison between early bilinguals and late learners suggests that early language experience gives early bilinguals advantages in the phonetic domain, even in cases in which the speakers have lost contact with their first language during childhood. This is the case of the bilingual Korean-Swedish participants studied by Hyltenstam, Bylund, Abrahamsson and Park (2009), who were adopted in childhood by Swedish-speaking couples and therefore lost exposure to their L1 Korean until adulthood. Compared to Swedish L2ers of Korean, who had no previous contact with Korean, the adoptees performed similarly in tasks requiring morpho-syntactic knowledge, but scored better in speech perception tasks. Furthermore, Au, Knightly, Jun, and Oh (2002) and Knightly, Jun, Oh, and Au (2003) compared L2ers and HSs with passive knowledge of the target language (Spanish or Korean) in both domains and concluded that HSs' pronunciation is more native-like than that of L2ers, even if they only had restricted contact with their HL during early childhood. Moreover, Flege, Yeni-Komshian and Liu (1999) claim that age affects the acquisition of phonology to a greater extent than morpho-syntax. The authors advance several explanations that account for this asymmetry, such as the varied timing of acquisition of those linguistic domains or the distinct degrees of similarity between L1 and L2 phonological and morpho-syntactic systems (see Flege, Munro, & McKay, 1995, for further discussion). Furthermore, other studies that focus on age constraints in language acquisition show that early acquirers outperform late learners in phonetic tasks both at the production and perception levels, strengthening the role of the age factor (Flege et al., 1995; MacKay, Flege, Piske, & Schirru, 2001; Piske, Flege, MacKay, & Meador, 2002; Flege & MacKay, 2004).

On the basis of these findings, we formulate the second hypothesis.

Hypothesis 2: If the HSs' knowledge is mainly shaped by aural input from birth, heritage bilinguals will differ significantly from L2ers in their pronunciation, and will be perceived predominantly as native speakers of EP.

An additional factor that is closely related with age and reduced language experience is cross-linguistic influence. Both L2ers and HSs have another linguistic system that co-exists with that of the target-language, either the L1 in the case of the L2 speakers or a dominant L2 (or also L1) in the case of the heritage bilinguals. Many authors have argued that cross-linguistic influence is a natural outcome of the contact between two languages in different bilingual acquisition settings (Jarvis & Pavlenko, 2008), even though there is still an intense debate about the nature and extent of its impact on early and late bilinguals and, specifically, on HL acquisition (see a recent discussion in Albirini & Benmamoun, 2014). Several studies have interpreted the HSs' performance as a consequence of dominant language transfer that resembles L1 transfer effects in L2 acquisition (e.g., Cuza & Frank, 2015). However, the apparent similarities between heritage and L2 speakers have to be interpreted with caution, since commonalities may be only superficial, resulting from a variety of sources. In particular, the HSs' low results in certain linguistic tasks may reflect a tendency to amplify variation that is present in the input and may not be the effect of dominant language transfer, as suggested by Rinke and Flores (2014).

3. STUDY 1: MORPHO-SYNTACTIC KNOWLEDGE

This section presents a follow-up study of Rinke and Flores (2014) that compared monolingual speakers and HHs of EP in their morpho-syntactic knowledge of clitics. Using the same method as the 2014 study, an offline AJT, this study includes a third experimental group, namely L2ers of EP whose L1 is German.

3.1 Participants

A total of 48 adult participants were included in the present study: (i) 16 second-generation HSs, who grew up in Germany; (ii) 16 German upper-intermediate L2 speakers of EP, who started to learn Portuguese in post-puberty age; and (iii) a group of 16 monolingually raised native speakers of EP.

At the moment of data collection, the HSs were between 20 and 43 years old (mean age = 28.8 yrs.; standard deviation [SD] = 6.7 yrs.). Eleven participants were born in Germany, while the other five immigrated in early childhood (until the age of five). All had parents who spoke mainly Portuguese at home. They were exposed to Portuguese from birth and to German before the age of five. Background information was collected in an oral interview and a written questionnaire, focusing on biographical information, language use and language preference. The questionnaire included a self-assessment scale, ranging from 1 (poor) to 4 (very good), regarding the following skills: oral and written comprehension and production. In their answers to questionnaire, all speakers claimed to use German more frequently in their daily life and to feel much more comfortable speaking the majority language. The use of Portuguese was reported to be mainly restricted to communication within the family, although at the time of testing it was no longer the exclusive language used within familial contexts, in which German also played an important role. Nonetheless, no speaker rated his/her proficiency in Portuguese to be lower than 3 (good), except in their writing skills, in which ten participants indicated intermediate proficiency. Contrary to the typical profile described in studies with Spanish HSs in the U.S. (Montrul, 2010; Benmamoun, Montrul, & Polinsky, 2013), all HSs included in our experiment had some formal instruction in their heritage language. All participants attended an extracurricular HL class for 6 to 12 years (mean years of attendance = 8.8 yrs.; SD = 2.3 yrs.) simultaneously with the regular German curriculum. This course on Portuguese language and culture consisted of one two-hour lesson per week, which mainly focused on developing basic reading and writing skills and some knowledge of Portuguese history and culture. At the time of testing, no participant was attending such a course, having concluded it five to 25 years before. The contact with written sources of Portuguese was almost non-existent, since they reported that they did not read Portuguese literature or newspapers and rarely visited Portuguese sites on the internet. Thus, exposure to EP was mainly restricted to aural input in colloquial and informal settings (for a more complete description of the profile of EP HSs, see Flores, 2015).

The group of L2 speakers included 16 participants with ages ranging from 22 to 36 years (mean age = 25.3 yrs.; SD = 3.6 yrs.). All participants were native speakers of German who started to learn Portuguese at the university, where the European variety is taught. Importantly, all L2ers were studying the language at the undergraduate or graduate level, attending language, linguistic and literature courses in EP. Background information was also gathered on the basis of a written questionnaire, which focused on AoA, amount and

type of contact with EP, language use habits and knowledge of other languages. They had been learning EP for at least two years and had successfully passed the B1.2 proficiency test of the Common European Framework for Languages (CEFR) for Portuguese, which was used as an assessment tool for the present study. In all cases the main source of contact with EP was the classroom setting, where several teaching methods were combined (e.g., explicit explanation of grammar; guided conversations; work with authentic and didactic materials; reading literature). Most participants had visited Portugal at least once. Three participants were at an Erasmus interchange program in Portugal at the time of testing. Besides the visits to Portugal and little contact with Portuguese friends, music and films, language input was mainly restricted to the classroom context and to contact with written sources of EP, for instance, through Portuguese literature, classroom textbooks and Portuguese websites.

In sum, the characteristic that both speaker groups had in common was their dominance in German. The groups differed with respect to the acquisition setting, the amount and type of EP input, and the AOA. While the heritage bilinguals had acquired EP in a naturalistic setting during childhood, the L2 speakers learned the language mainly through explicit language instruction at a post-puberty age. Even though their contact with EP was limited compared to monolinguals, HSs still had more input than L2ers, since Portuguese was one of their home languages.

The monolingual group comprised 16 native speakers of EP from 17 to 49 years of age (mean age = 28.3 yrs.; SD = 11.1 yrs.). No participant had lived abroad or had been raised bilingually. Regarding their level of education, 11 informants had a basic degree (around 9 years of schooling) and five had a high school degree (12 years of schooling). All the monolingual speakers were born and raised in Portugal.

3.2 The EP Clitic System, Test Conditions and Hypotheses

The clitic system of European Portuguese is of special interest because of its variability and complexity and because it includes aspects stabilized through the formal and written register, as well as properties that vary between the vernacular and a more formal register.

EP has two types of object pronouns that are traditionally described in terms of an accented and unaccented paradigm. In Cardinaletti and Starke's (1999) terminology, one can speak about a paradigm of strong and clitic pronominal forms. The strong forms can be emphasized and are independent syntactic constituents, whereas the clitic forms behave syntactically as heads, cannot be stressed and are attached to a verbal host from which they cannot be separated (examples 1a.-b.).

- (1) a. A Maria encontrou-me.
the Mary met me_{ACC} (neutral)
'Mary met me.'
- b. A Maria encontrou-me a MIM e não a ele.
the Mary met ME and not him (emphasis on the direct object)
'Mary met me and not him.'

- c. (O Francisco não estudou para o teste de matemática.)
the Francisco not studied for the test of mathematics
- O professor chumbou **ele*. [correct: chumbou-o]
the teacher failed him_{strongACC}
'Francisco did not learn for the math test. The teacher failed him.'
- d. (O Luís trouxe os seus amigos para a sua casa para jogarem Playstation.)
the Luis took the his friends to the his home to play Playstation
??A mãe ofereceu uma pizza a *eles*.
the mother offered a pizza to them_{DAT}
'Luis took his friends home to play Playstation. His mother offered pizza to them.'

As shown in 1b to 1d, the use of the strong form requires the parallel realization of the clitic form, even though in colloquial EP the sentence is less deviant if the strong pronoun is a dative pronoun as in 1d. (see footnote in Brito, 2008).

Given that Standard German freely allows for the occurrence of isolated strong pronouns in object position and does not have EP-like clitics, the acceptance of strong pronouns as in 1c or 1d may be a reflex of cross-linguistic influence because speakers may transfer the German option to Portuguese. However, transfer is only probable if the speakers are not sensitive to the dative-accusative asymmetry (1c. versus 1d.). If speakers show sensitivity to the case asymmetry, i.e., if they reject the examples in the ungrammatical dative condition to a lesser extent than in the accusative condition, cross-linguistic influence is less likely. In addition, this reflects intuitive native knowledge that cannot be acquired on the basis of instruction or formal input, because both constructions are ungrammatical in the standard language.

EP clitics show a further number of peculiarities in comparison to other Romance languages. They vary concerning their position depending on the syntactic context and in their morphological form depending on the phonological context (2a.-c.).

- (2) a. A Maria viu-**o**. (canonical form)
the Mary saw him_{cliticACC}
'Mary saw him.'
- b. As meninas viram-**no**. (allomorph, following nasal ending of the verb)
the girls saw him_{cliticACC}
'The girls saw him.'
- c. A Maria pode **fazê-lo**. (allomorph, *o* becomes *lo* after *-r*, which drops).
the Mary can do it_{cliticACC}
'Mary can do it.'

The allomorphs of the clitic pronoun are quantitatively rare compared to the canonical form (Flores, Rinke, & Azevedo, 2017). They are explicitly taught in school because many EP children have difficulties with the allomorphic forms. If speakers show

deviances in the allomorphic conditions, this may reflect less exposure to formal or written input sources.

In contrast to most other Romance languages, EP allows for Romance-type as well as English-type topicalization, that is, topicalization with and without clitic resumption (Duarte, 1987):

- (3) a. A Linda não gosta de doces, mas o teu bolo ela comeu__{/(-O_{ACC})}
the Linda not likes *prep* sweets but the your cake she ate
com muito gosto.
with much pleasure
'Linda doesn't like sweets but she ate your cake with great pleasure.'
- b. À sua mãe, o João ofereceu__{/(-lhe_{DAT})} flores.
To his mother the João offered (her) flowers
'To his mother, John offered flowers.'

As shown by the syllabi of EP L2 courses, topicalization structures (with clitic resumption) are to some extent part of L2 instruction. However, the variability of resumption constitutes a language-internal asymmetry that is not explicitly addressed in L2 courses and HL classes. In addition, topicalization seems to include a grammar-internal case asymmetry: monolingual EP speakers seem to prefer topicalization of dative objects without a clitic (3b.) over topicalization of accusative objects without resumption (3a.).

A variety of outcomes are possible with respect to topicalization structures: Given that German does not show resumption in topicalization structures, HSs and L2ers may reject topicalization with resumption as an outcome of cross-linguistic influence. If the speakers rely on instruction, they may reject topicalization without resumption because this structure is not explicitly taught. Additionally, native competence implies that speakers are able to recognize the variability of resumption and are sensitive to the case asymmetry. Variability also arises with respect to clitic climbing, that is, in complex verb constructions, it can either be attached to the non-finite verb or 'climb up' to attach to the finite verb, as shown in (4).

- (4) a. O João *pode-o* comprar na semana que vem.
the John can-it_{clitic} buy in-the week that comes
- b. O João pode *comprá-lo* na semana que vem.
the John can buy-it_{clitic} in-the week that comes
'John can buy it next week.'

The non-climbing option (4b) seems to be preferred in formal registers, whereas climbing (4a) is a typical and frequent feature of spoken EP (see Andrade, 2010). If speakers rely on knowledge acquired on the basis of instruction or formal input sources, they may tend to reject the climbed option (4a.). Reliance on spoken input sources is reflected by the acceptance of both options to the same degree. Given that pronominal objects occur to the left of the lower verb in German in complex verb structures, rejection of the non-climbed option (4b.) could also reflect cross-linguistic influence.

Taking the research hypotheses into consideration, several aspects of the Portuguese clitic system mentioned so far can be related to registers and input sources. We expect monolingual speakers to reflect, to some extent, variation present in the spoken language and to know properties relating to more formal input. However, given that this test is written and untimed, speakers' judgments can be biased towards the formal register. HSs may perform more native-like than L2ers with respect to the variation found in topicalization structures and clitic climbing because knowledge of those properties relies on spoken input. Concerning allomorphic clitic forms, L2ers may instead have advantages. Transfer effects both among L2ers and HSs may be shown in a tendency to reject the non-climbed clitic order, to accept, equally, dative and accusative strong pronouns in isolation, and to reject topicalization with resumption.

3.3 Methodology

For this experiment we replicated a reduced version of an untimed written Acceptability Judgment Task used by Rinke and Flores (2014). The replication was restricted to the four condition types mentioned in the previous section: a) strong pronouns in isolation, b) allomorphy of clitic forms, c) topicalization with and without resumption, and d) clitic climbing. The conditions divided into 12 subconditions with five tokens each (a total of 60 items). Four sub-conditions refer to ungrammatical and eight sub-conditions to grammatical EP sentences. The sentences were randomized and the participants were asked to judge whether the sentence "sounds correct" or "sounds incorrect." Additionally, participants were asked to correct all sentences they judged as "sounding incorrect" in order to find out their motivation for rejecting the examples. This type of task might increase the likelihood of accessing explicit grammatical knowledge in domains fostered by formal instruction but it also reveals knowledge of variable domains and implicit asymmetries, since speakers were required to judge structures that are frequent in the spoken language. The test was performed in the presence of a researcher and took approximately 30 minutes to complete.

For the statistical analysis the accuracy rate of the answers was quantified. In the case of ungrammatical conditions, accuracy refers to the rejection and accurate correction of the ungrammatical items. With respect to grammatical conditions, accuracy refers to the acceptance of the items. Rejection of grammatical items and their correction using other structures indicates a preference for alternative constructions. The statistical analysis was performed on IBM SPSS Statistics 21.0. Since the data were not normally distributed, non-parametrical tests were performed for the statistical analyses.

3.4 Results

3.4.1 Overall Distribution

Table 1 gives an overview of the overall distribution of accuracy scores (%) per group.

Table 1.*Mean Accuracy, Standard Deviation, Minimum and Maximum per Group (Study 1)*

Group	N	Mean	Std. Deviation	Min.	Max.
Monolingual Speakers	16	88.1	5.8	76.0	98.0
Heritage Speakers	16	57.2	12.9	38.0	84.0
L2 Speakers	16	64.5	11.0	38.0	88.0

The mean level of accuracy was about 88.1% (SD = 5.8) in the monolingual group, contrasting with the significantly lower rates of the other two groups. The HSs reached a mean rate of 57.2% (SD = 12.9) of accurate responses, while the mean average of accuracy was about 64.5% (SD = 11.0) in the L2 group. Additionally, the higher standard deviation and the greater distance between minimal and maximal scores in the latter groups demonstrate that both heritage and L2 speakers also showed significantly higher individual variation.

A Kruskal Wallis test confirmed that the three groups differed significantly ($H(2) = 30.135, p < .001$). Two additional Mann-Whitney tests indicated that both the heritage and the L2 speakers differed significantly from the monolingual controls (HS versus MS: $U = 5.000, p < .001$; L2 versus MS: $U = 7.500, p < .001$).¹ In contrast, the heritage bilinguals did not differ from the L2 group concerning overall accuracy scores ($U = 78.00, p = .059$). As expected, both L2 learners' and heritage bilinguals' responses deviated from the monolingual control group's. In the following section, we will describe the results per condition.

3.4.2 Pronoun Choice: Strong Pronouns Instead of Clitics

Table 2 shows the results of the speakers' intuitions concerning the use of strong pronouns instead of clitics.

Table 2.*Pronoun Choice: Strong versus Clitic Pronoun, Mean Accuracy and Standard Deviation per Group (Study 1).*

Linguistic property	Condition	Monolingual Speakers	Heritage Speakers	L2 Speakers
Choice (strong versus clitic pronouns)	a) Use of strong dative pronoun instead of clitic (ungrammatical)	93.8 (9.6)	27.1 (34.8)	55 (35.4)
	b) Use of strong accusative pronoun instead of clitic (ungrammatical)	100 (0)	54.4 (36.9)	58.8 (34.6)

The results showed that monolingual speakers did not accept the use of strong pronouns instead of accusative clitics (100% accuracy), but they showed some variation regarding the use of strong dative pronouns (93.8% accuracy, $SD = 9.6$). The heritage bilinguals also accepted the ungrammatical use of strong dative pronouns significantly more than strong accusative pronouns (27.1% of accuracy in dative and 54.4% in accusative contexts), although their performance was low in this condition. In contrast, L2 speakers showed very similar results in both sub-conditions (55.0% of accuracy with strong dative pronouns, $SD = 35.4$, and 58.8% strong accusative pronouns, $SD = 34.6$). A Wilcoxon test confirmed that only in the L2 group the difference between both sub-conditions was not significant ($Z = -.354, p = .723$).

3.4.3 Allomorphy of Clitic Forms

The condition testing allomorphy of clitic forms included four sub-conditions; two were grammatical and two ungrammatical. In Table 3, the mean average of accuracy per group is organized per sub-condition.

Table 3.

Allomorphy of Clitic Forms, Mean Accuracy and Standard Deviation per Sub-Condition and Group (Study 1)

Linguistic property	Condition	Monolingual Speakers	Heritage Speakers	L2 Speakers
Allomorphy of clitic forms	Use of <i>-no/-na</i> (grammatical)	99.9 (0.3)	56.4 (24.5)	83.8 (18.2)
	Use of <i>-o/-a</i> instead of <i>-no/-na</i> (ungrammatical)	80.0 (30.1)	7.5 (21.8)	55.0 (37.6)
	Use of <i>-lo/-la</i> (grammatical)	95.0 (11.6)	76.3 (24.5)	91.3 (16.3)
	Use of <i>-o/-a</i> instead of <i>-lo/-la</i> (ungrammatical)	97.5 (6.8)	51.3 (41.3)	65.0 (37.6)

The first interesting observation is related to the HSs' performance, which had considerably lower scores of mean accuracy than the L2 speakers in the four conditions. With respect to the use of the allomorphs *-no/-na*, both L2ers and HSs also differed significantly from the monolinguals, who reached considerably higher accuracy scores (grammatical condition: $H(2) = 28.560, p < .001$; ungrammatical condition: $H(2) = 26.147, p < .001$). Nevertheless, the performance of the HSs contrasted significantly with the average of accuracy of the L2 group: the HSs only reached 56.4% of accurate responses ($SD = 24.5$) in the grammatical condition of *-no/-na* and 7.5% in the ungrammatical condition ($SD = 21.8$). This means that they did not recognize the grammatical use of these forms in many contexts and were not aware of the ungrammatical use of *-o/-a* in contexts where *-no/-na* is required. The L2ers' answers had a 83.8% ($SD = 18.2$) accuracy rate with the grammatical condition of *-no/-na* and a 55.0% ($SD = 37.6$) accuracy rate with the ungrammatical condition. Two Mann-Whitney tests confirmed that the difference between both groups was statistically significant in both conditions (grammatical condition: $U = 51.500; p < .001$; ungrammatical condition: $U = 49.500; p < .001$).

Considering the use of the allomorphic forms *-lo/-la*, both in the grammatical as well as in the ungrammatical condition, the results of all groups of speakers are much higher than in the case of *-no/-na*, especially in the monolingual and in the HS groups. Although the three groups also differed statistically in both *-lo/-la* conditions, as shown by two Kruskal-Wallis tests (grammatical condition: $H(2) = 7.953, p = .019$; ungrammatical condition: $H(2) = 15.161, p = .001$), there was no statistical difference between the heritage and the L2 speakers regarding this particular structure ($U = 91.000; p = .070; U = 119.500; p = .384$, respectively).

Interestingly, monolingual speakers showed significant differences concerning their judgments of ungrammatical *-no/-na* versus *-lo/-la*. In particular, monolinguals showed some tendency to accept *-o/-a* in contexts of *-no/-na* but not in contexts of *-lo/-la*. This allomorphic difference was statistically significant ($Z = -2.585, p = .010$). Like the monolingual speakers, the heritage bilinguals also showed significant differences between the ungrammatical use of *-no/-na* and *-lo/-la* ($Z = -3.000, p = .003$). Conversely, L2ers seem not to be sensitive to this allomorphic difference ($Z = -1.035, p = .301$). To conclude, HSs performed less accurately than L2ers in this condition. Nevertheless, like monolingual speakers and unlike L2ers, they are sensitive to language-internal asymmetries.

3.4.4 Topicalization

Table 4 refers to the speakers' intuition of topicalization structures with and without a clitic pronoun, both in accusative and dative contexts.

Table 4.

Topicalization Structures, Mean Accuracy and Standard Deviation per Group (Study 1)

Linguistic Property	Condition	Monolingual Speakers	Heritage Speakers	L2 Speakers
Topicalization structures without a clitic	a) Topicalization without an accusative clitic (grammatical)	65.0 (26.8)	60 (29.2)	51.9 (32.5)
	b) Topicalization without a dative clitic (grammatical)	87.6 (17.7)	81.3 (25.8)	48.1 (42.5)
Topicalization structures with a clitic	c) Topicalization with an accusative clitic (grammatical)	98.8 (6.5)	88.8 (16.8)	77.5 (25.2)
	d) Topicalization with a dative clitic (grammatical)	96.3 (24.9)	87.5 (22.5)	78.1 (23.8)

The results indicate that both monolingual and HSs showed sensitivity to case in topicalization structures without resumption. Topicalization of dative objects without a clitic was much more frequently accepted than topicalization of accusative objects without resumption. This mismatch between both sub-conditions was statistically significant in both groups, as shown by the results of a Wilcoxon test (monolinguals: $Z =$

-2.944, $p = .003$; HSs: $Z = -2.626$, $p = .009$). Additionally, two Mann-Whitney tests confirmed that HSs did not differ statistically from monolingual speakers in neither sub-conditions (accusative: $U = 144.50$, $p = .571$; dative: $U = 129.50$, $p = .257$).

In contrast to the heritage and monolingual speakers, L2 speakers did not show sensitivity to case distinction in topicalization structures. They performed similarly in both sub-conditions (accusative: 51.9% of acceptance, $SD = 32.5$; dative: 48.1% of acceptance, $SD = 42.5$). A Wilcoxon test confirmed that the difference between both sub-conditions was not significant in the L2 group ($Z = -.660$, $p = .509$). This means that L2 speakers showed low acceptance of topicalization structures without a resumptive clitic both in the accusative and the dative condition. In these latter contexts, the three groups differed statistically, as indicated by a Kruskal-Wallis test ($H(2) = 9.564$, $p = .008$). The results in the conditions with resumption showed that topicalization with clitics (clitic left dislocation, Duarte, 1987) is more easily accepted by all speaker groups than topicalization without resumption, although both structures are grammatical. L2 learners are in principle able to use topicalization structures although they hardly recognize that topicalization without clitic resumption is part of the EP grammar. Unlike monolingual speakers and heritage bilinguals, they also do not differentiate between the dative and the accusative condition in topicalization structures without resumption.

3.4.5 Clitic Climbing

In this condition, the speakers' knowledge of clitic placement in restructuring contexts was tested in a) sentences in which the clitic precedes the main verb and b) the clitic is in post-verbal position.

Table 5.

Clitic Climbing, Mean Accuracy and Standard Deviation per Group (Study 1)

Linguistic property	Condition	Monolingual Speakers	Heritage Speakers	L2 Speakers
Clitic climbing	a) climbing (grammatical)	65.3 (29.2)	76.3 (32.0)	51.9 (30.3)
	b) non-climbing (grammatical)	96.3 (10.9)	83.4 (22.3)	85.0 (26.8)

The results demonstrate that, even though both clitic positions are grammatical, monolingual speakers showed a clear tendency to favor the post-verbal clitic position (96.3% acceptance) and to reject clitic climbing (65.3% acceptance) in this type of task. L2 speakers performed similarly to monolinguals, that is, they accepted clitic climbing only in 51.9% of all given sentences, while the rate of acceptance of post-verbal clitics was much higher (85.0%). Two Mann-Whitney tests confirmed that monolingual and L2ers did not differ statistically in these contexts (a: $U = 125.500$, $p = .528$; b: $U = 105.500$, $p = .187$). Additionally, two Wilcoxon tests demonstrated that the mismatch between pre- and post-verbal clitic placement was significant both in the monolingual ($Z = -3.154$, $p = .002$) and the L2 group ($Z = -2.463$, $p = .014$).

In opposition to the former groups, HSs did not favor the post-verbal clitic position. Their average of acceptance was similar in both sub-conditions (76.3%, $SD = 32.0$; 83.4%, SD

= 22.3). As expected, a Wilcoxon test confirmed that the difference of the heritage bilinguals' acceptance rate regarding pre- versus post-verbal clitic placement in restructuring contexts was not statistically significant ($Z = -.632, p = .527$).

3.5 Discussion

Firstly, the overall mean rate of accuracy of the HS group, which does not differ from the L2 speakers' results, corroborates previous findings that relate the results of heritage bilinguals in written AJTs to overall difficulties in written tests that require some metalinguistic knowledge (Bowles, 2011; Cuza & Frank, 2015). Overall, heritage and L2 speakers seem to be less confident of their intuitions than monolinguals and they show more individual variation. Despite these similarities, we observed important differences between both groups of speakers.

Our first observation concerns variability within the group of monolingual speakers, which reflects variation present in the spoken register of the Portuguese language. One example is the use of strong pronouns in object position. Spoken EP shows some variation concerning the use of strong dative pronouns instead of clitics (Brito, 2008), whereas this variation is not reported for accusative objects. The monolinguals' results reflect this asymmetry between dative and accusative pronouns. Interestingly, the same asymmetry is observable in the performance of the HSs, who are much more amenable to accept strong dative pronouns than strong accusative pronouns, although at a much lower level of accuracy. In contrast, L2 speakers did not distinguish between dative and accusative pronouns, presenting mean rates at chance level in both conditions. Given that German allows for strong pronouns in object position, cross-linguistic influence could be reflected on an overall erroneous acceptance of strong pronouns in object position by L2 and HSs. However, the asymmetry between accusative and dative, observed in the HSs' group, cannot be attributed to the influence of German. Therefore, a phenomenon that appears to be the outcome of transfer is, in fact, the outcome of variation also present in the performance of native speakers. Inversely, in the case of the L2ers, the absence of this asymmetry could, indeed, reflect transfer effects. However, the good performance of the L2 speakers in other conditions with clitic pronouns, e.g. topicalization with resumption, contradicts this idea. In fact, neither L2ers nor HSs showed a tendency to reject topicalization with clitic resumption as would have been expected under the influence of German.

Regarding topicalization structures, monolingual speakers showed a clear tendency to accept the omission of the clitic more readily when a dative object was preposed. HSs did not differ from the monolingual controls. They showed approximately the same rates of accuracy and were sensitive to the dative-accusative asymmetry. Conversely, the L2 speakers did not show sensitivity to case distinction in topicalization structures. They accepted these structures approximately at chance level in both cases. Their responses indicate that they did not have intuitions regarding this variation, which typically occurs in spoken registers. Conversely, HSs show implicit grammatical intuitions due to their dominant exposure to oral Portuguese from early on.

Contrary to the first two structures, allomorphic clitic forms are acquired late and consolidated through formal education or intensive contact with written registers. School teachers of Portuguese language report that pupils tend to fail in choosing the appropriate enclitic form until they reach high levels of schooling. In fact, this domain is the one in

which HSs perform worse than L2ers. This shows that, despite some instruction in HL programs, these courses, combined with absent contact with written Portuguese at time of testing, were not sufficient to stabilize the HSs' knowledge of this property. However, in this subcondition, monolingual speakers also showed an asymmetry between the forms *-no/-na* and *-lo/-la* that has to be explained through variation in oral input. It appears that the phonetic proximity between the nasal consonant in *-no/-na* ([nu] / [nɐ]) and the nasal ending of the preceding verb form (e.g. *comem-no*, [kɔmãj.nu] 'eat-them') contributes to the difficulties in judging allomorphs, in native speech. Interestingly, the asymmetry between both shaped allomorphic clitic forms is present in the monolinguals' and the HSs' judgments, but not in the L2 speakers' performance. This shows that, although heritage bilinguals have general difficulties with the allomorphic clitic forms of EP, they are, nevertheless, sensitive to the phonological realization of the two different forms, whereas L2 speakers lack this phonological sensitivity and do not mirror the asymmetry between the two forms in the EP system.

The last context testing the differing input conditions of HSs and L2ers included variable clitic placement in restructuring contexts. Clitic climbing with semi-auxiliaries such as *poder* ('can') is truly optional in EP. Nevertheless, in the AJT, monolingual EP speakers often corrected the grammatical examples with clitic climbing into the non-climbed structure. This performance seems to be related to a difference between oral and written registers, since the climbed option occurs more often in spoken language, whereas the non-climbed order is more frequent in formal and written language, as shown by Andrade (2010). Interestingly, clitic climbing is also a way to avoid the use of allomorphs, since in the non-climbing construction, the main verb appears in the infinitive, which ends with a vibrant consonant and triggers the use of *-lo/-la*. Climbing avoids the allomorph in these constructions (see Sentence 2c, which would become *A Maria pode-o fazer*), and could be an avoidance strategy for HSs, who struggle with allomorphs. L2ers, who presumably base their judgments on explicit knowledge and have more contact with written texts than HSs, show the same tendency as monolinguals to reject grammatical clitic climbing. The HSs, who have less contact with the formal written register, performed differently from the other two groups in this condition and accepted the clitic climbing structures and the non-climbed option at a similar level. Again, transfer does not seem to play a role in this condition, because heritage bilinguals do not show a tendency to reject the word order variant, which is ungrammatical in German (non-climbing). The L2-learners even accept this word order variant to a greater degree than the climbed order, which would correspond to the German word order.

Overall, the reliance of HSs on oral input sources is shown by their better performance in conditions that are related to the oral register. L2 learners instead seem to rely on written/formal sources.

The test also reveals that there are some aspects of the grammar that are related to intuitive knowledge about inner-grammatical variation: that strong dative pronouns are somehow less deviant than strong accusative pronouns, that topicalization of datives without resumption is more easily acceptable than accusative objects without resumption in topic position, and that after a nasal, the initial nasal of the allomorph *-na/-no* may more easily be omitted than the initial lateral of the allomorph *-la/-lo* in intervocalic position. This knowledge is shared by monolingual speakers and heritage bilinguals, but not by L2 learners and can be either attributed to the relevance of spoken input sources for the

acquisition of these variable properties or to different AoA in the HSs and L2 learner group. These results thus confirm our first research hypothesis.

Finally, we discuss the results of the intra-group variation. As shown in Table 1, standard deviation is significantly higher in the HS and L2 groups, which mirrors high variability within the groups. A factor that might contribute to this variation is the proficiency level of the L2ers or the years that the HSs received heritage language instruction. In order to evaluate the relationship between the overall accuracy rate and the L2ers' proficiency level, a Spearman correlation test was applied. This test showed no correlation between the variables CEFR level and overall outcomes of the participants ($r = .405$, $p = .120$). Curiously, the participant with the highest rate of accuracy (L2_3: 88%) had only an intermediate (B1+) proficiency level.

In contrast, a correlation test to verify the association between the rate of accuracy and the HSs' length of HL instruction (in this case, we performed a Pearson Correlation Coefficient test, since both variable are ordinal) showed a significant positive effect ($r = .652$, $p = .006$). This means that the mean rate of accuracy increases with increasing years of attending a HL course. This adds to the conclusion that the heritage bilinguals' I-grammar is mainly shaped by naturalistic exposure to spoken EP and reflects variability present in the oral input. Exposure to formal registers and instruction seems to sharpen the speakers' intuitions and stabilize knowledge of certain properties, since speakers with more years of attendance of HL courses show higher rates of accuracy in the AJT, namely in the allomorphic condition. On the contrary, a higher level of proficiency is not related to higher accuracy scores in the L2 group, which may be explained by the conditions included in the test. As previously mentioned, most properties are shaped by oral input, but the differing proficiency levels may not reflect a significant difference in input quantity.

4. STUDY 2: GLOBAL ACCENT

Study 2 consisted in a global accent rating task in which speech samples of three speaker groups of EP, namely heritage bilinguals, L2 learners and monolingual speakers, were rated by monolingual EP listeners according to their degree of nativeness.

4.1 Participants

Two types of informants participated in this experiment: 24 speakers of European Portuguese, whose speech samples were assessed in the global accent rating task, and 46 Portuguese listeners, who performed as raters. The cohort of 24 EP speakers consisted of three groups: (i) 12 Lusophone HSs (10 female (F) and 2 male (M)) from 19 to 30 years of age (mean age = 23.08 yrs; SD = 4.01 yrs), who grew up in Germany or Switzerland; (ii) 6 German learners of L2 EP (4F and 2M) between 25 and 42 years of age (mean age = 30.33 yrs; SD = 6.53 yrs), and with an upper-intermediate level of L2 proficiency; and (iii) 6 monolingual Portuguese speakers between 26 and 43 years of age (mean age = 32.17 yrs; SD = 7.65 yrs). The group of speakers in this study was smaller than in Study 1 due to the need to select homogeneous speech samples, which fulfilled the criteria described in Section 4.2 in terms of content and linguistic structure, and to avoid a lengthy global accent task. Therefore, only six HSs participated in both studies, since some sound files had to be excluded for technical reasons. However, all HSs were chosen on the basis of the same criteria and shared a similar background in terms of AoA of EP and German, family structure, language use, and attendance of HL courses, and thus had exactly the

same speaker profile as the informants described in Study 1. The 12 HSs grew up in Germany/Switzerland with Portuguese as their home language and German as the preferred dominant language. All were raised in first-generation families and attended the HL course for 7 to 12 years (mean = 10.0 yrs; SD = 1.76). For the L2 group, we selected proficient L1 German speakers on an upper intermediate (B2 CEFR) level and used EP in their daily life for professional purposes. They were either studying Portuguese in a graduate program at the university or had already concluded their bachelor's degree and were using EP in their work. All learned EP in a formal setting, but also contacted with EP outside the classroom through frequent visits to Portugal or close relationships with Portuguese friends.

The 46 monolingual EP listeners (28 F and 18 M) that formed the group of raters were all undergraduate students, with ages between 18 and 38 years (mean age = 20.61 yrs; DP = 3.06 yrs), and were born and raised in Portugal. Students from the undergraduate course of modern European languages and literatures with basic knowledge of and training in (articulatory) phonetics were selected because experience with linguistics and other languages can have a positive effect on inter-rater reliability (Thompson, 1991, as cited in Jesney, 2004). All raters reported having no hearing impairments.

4.2 Procedure and Research Question

The stimuli included in the global accent rating task were speech samples collected in naturalistic semi-spontaneous production tasks (biographic interview, story narration and picture description). The speech samples from the monolingual speakers were extracted from a spoken corpus of interviews available at the University of Minho.² Three sentences from each of the 24 speakers, totalling 72, were selected based on the following criteria: Samples consisting of (i) full sentences with (ii) no morphological, syntactic or lexical deviations, (iii) no hesitations, (iv) no long pauses (<0.3 sec), and (v) no cultural information. Although the duration of the sentences was within the limits of auditory sensory memory (i.e., echoic memory), which is about 2 to 4 seconds (Goldstein, 2011), the three sentences were concatenated with a two-second inter-stimulus interval (ISI) to provide sufficient segmental and suprasegmental information for the ratings. In terms of mean duration, the sentences produced by the three groups of speakers did not differ significantly ($F(2,21) = .37, p > .05$), ranging from 3.5s to 3.7s (mean duration = 3.57s; SD = 1.5s). Table 6 presents information on the mean duration of sentences.

Table 6.

Mean Duration of Sentences

Speakers	Mon	HS	L2
Mean duration (SD)	3.6s (0.4s)	3.5s (0.6s)	3.7s (0.3s)

The 24 stimuli (sequence of 3 sentences x 24 speakers) were presented twice in randomized order. Since previous research has provided evidence that familiarization with the stimuli set tends to lead to harsher judgments (e.g., Flege & Fletcher, 1992; Munro & Derwing, 1994), only the ratings of the second presentations were analysed in order to have more reliable ratings.

Stimuli were presented to the listeners individually over headphones with a software for speech perception experiments, *TP (Perceptual Testing/Training) v.3.1.* (Rauber, Rato, Kluge, & Santos, 2012). Before the test, instructions (which included the indication that regional accents should be identified as native) were first explained orally and then shown in written form in the test screen. After this, listeners performed a training task with presentation of four speech sequences not included in the test stimuli: one from a monolingual L1 speaker, one from an L2 speaker with a clearly detectable foreign accent and two from bilingual speakers, so that the listeners could 'calibrate' their perception by identifying the two endpoints of nativeness.

The global accent rating task consisted of two steps, following the procedure of De Leeuw, Schmid and Mennen (2010) and Moyer (2004) and replicated by Hopp and Schmid (2013) and Kupisch et al. (2014). In the first step, after listening to a triad of sentences, informants made a binary judgment by labelling it as a sample of native EP speech or non-native speech. This choice was followed by the indication of degree of certainty in a three-point scale (1 = certain, 2 = semi-certain, 3 = uncertain). For statistical analysis, this scale was converted into a 6-point Likert scale by combining the two judgments (1 = certain of native speech, 2 = semi-certain of native speech, 3 = uncertain of native speech, 4 = uncertain of non-native speech, 5 = semi-certain of non-native speech; 6 = certain of non-native speech). Specifically, this task aimed at determining whether native speakers of European Portuguese perceived the accent of Portuguese heritage bilinguals who lived in a German-speaking country as (i) similar to the accent of Portuguese monolingual speakers or (ii) more similar to the accent of highly proficient German speakers of Portuguese L2.

After the task, the 46 raters answered a questionnaire about the most relevant factors (e.g., pronunciation of phonetic segments, word stress, or sentence prosody) determining the classification of stimuli as non-native speech samples and the identification of the native language of the speakers identified as non-native.

4.3 Results

For the present purpose, we will focus on the results of the 6-point rating scale that combines the binary choice ('native' – 'non native') with the 'degree of certainty.' Table 7 shows the mean ratings (ranging from 1 'certain of native speech' to 6 'certain of non-native speech') by group of speakers.

Table 7.

Global Accent Ratings (Mean and Standard Deviation per Group)

Global accent ratings	Monolingual Speakers	Heritage Speakers	L2 Speakers
Mean (SD)	1.07 (0.06)	1.66 (0.85)	5.82 (0.12)

Figure 1 shows the mean ratings received by each participant.

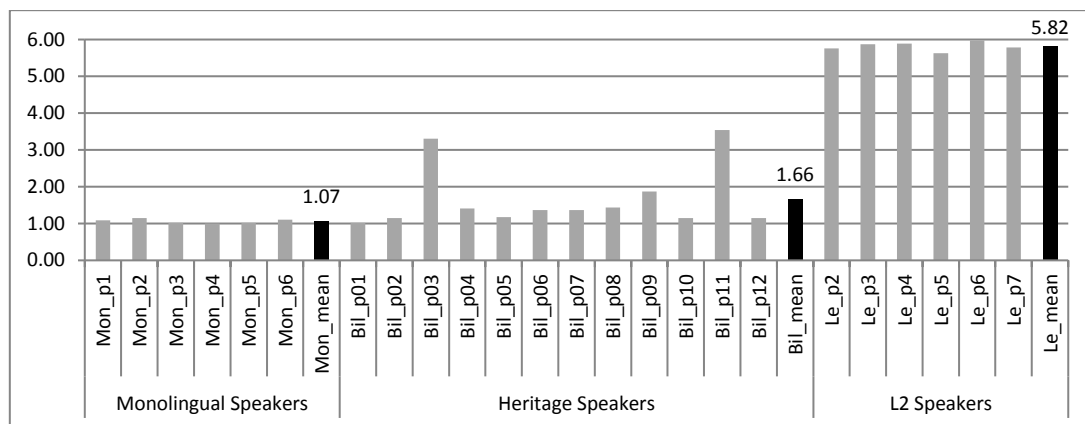


Figure 1. Degree of Nativeness per Group

The first relevant observation is related to the difference between the monolinguals and the L2ers that represented the endpoints of nativeness. In the monolingual group, the participants were consistently rated as native speakers of EP, with a high level of certainty. The mean rating was 1.07 (SD = 0.06), ranging from 1.00 to 1.15. By contrast, the L2 speakers were consistently rated as having a non-native accent. The mean rating was 5.82 (SD = 0.12) in this group, reaching from 5.63 to 6.00 average points.

The ratings of the heritage bilinguals showed more variation concerning the degree of nativeness of their accent in comparison to the monolinguals, which is a result reported by other studies (Hopp & Schmid, 2013). A majority of speakers (9 participants) performed very close to 1.00 point ('certainly a native speaker'), the average rating attributed to the monolinguals. One participant presented a mean rating of 1.87, a value that was still very close to the monolinguals' average (the bilingual speaker was rated as native speaker, but in some cases with a semi-certain degree of confidence). Two participants also showed more distant ratings compared to the monolinguals' scores (3.30 and 3.54, respectively). Their ratings were between the mean scores of the monolingual and the L2ers group. This suggests that the raters were less confident of their evaluation, indicating that the accent of these two HSs may have shown some particularities that set them off. However, their ratings were still distant from the L2 speakers' evaluation. Furthermore, even with these higher ratings, the mean score of the HSs' group was very close to the monolinguals' (mean = 1.66; SD = 0.85).

For the statistical analysis, non-parametric tests were used due to the unbalanced number of participants across groups and the small size of the sample. As expected, a Kruskal-Wallis test corroborated that the three groups differed significantly concerning the evaluation of their accent ($H(2) = 17.727, p < .001$). Follow-up Mann-Whitney tests with Bonferroni correction confirmed that monolinguals and L2ers differed significantly ($U = 0.000, p < .017$) and heritage bilinguals and L2 speakers did as well ($U = 0.000, p = .001$) in terms of their accent ratings. Due to the higher variation within the HS group, another Mann-Whitney test was performed, and the comparison between heritage bilinguals and monolingual speakers revealed significant differences between the groups ($U = 5.500, p = .004$).

4.4 Discussion

This experiment showed that HSs of EP are clearly distinct from proficient L2ers concerning pronunciation. Although with less certainty than native monolinguals, the majority of bilinguals were rated as having a native EP accent, thus confirming the claim that early exposure to a language, even in a context of minority language acquisition, is a strong predictor of native-like proficiency in speech production. Conversely, the late L2ers seem to have difficulties in attaining native-like pronunciation, even at high levels of proficiency. This is in line with a large body of studies that show the malleability of production (and perception) of phonetic categories in speakers who are beyond the critical period for language acquisition (Flege, 1995; Hyldenstam & Abrahamsson, 2003). However, even studies that attest small differences between early/simultaneous bilinguals and monolingual speakers (e.g., Guion, 2003; Mack, 1989 for adult grammars; Lleó, 2006; Paradis, 2001, for children) suggest that the former have phonological systems that are *largely*, though *not entirely*, independent in the two languages. There is indeed considerable evidence in favor of the idea that the phonological systems of bilinguals interact, leading to particularities in production (Fowler, Sramko, Ostry, Rowland, & Hallé, 2008; Stangen, Kupisch, Ergün Proietti, & Zielke, 2015) and that this interaction may extend over a lifetime (Godson, 2004). Unlike the studies that show deviant phonetic features in HSs' speech, our results attest a largely independent, native-like development of EP HSs' pronunciation. This difference may be explained by the divergence of EP HSs from other groups of heritage bilinguals with respect to the input they received in early years. The EP participants tested in this study were all raised in first-generation families whose clear dominant (or even only) home language was Portuguese. This supports the idea that the amount and type of input in early years are significant predictors of language development. The two 'divergent' bilinguals in the present study seem to reinforce the assumption that an overall independent development of phonology may nevertheless show singularities not observed in native speech. For the purpose of this paper, it is important to highlight that the HSs' phonetic performance in their heritage language is not equivalent to that of the L2ers, confirming our second research hypothesis.

5. CONCLUSION

Studies comparing L2 learners and heritage bilinguals have revealed similarities that distinguish the two groups from monolingual speakers and suggest that heritage bilinguals' linguistic competence is somehow divergent when compared with monolinguals (Montrul, 2011). On the other hand, it has been argued that heritage bilinguals have an advantage over L2 learners in oral tests and tasks that measure implicit rather than metalinguistic knowledge (Bowles, 2011). Our results indicate that the similarities between L2ers and HSs are only superficial and that the differences reflect fundamental characteristics that set both types of speakers apart. This was shown on the basis of two independent studies comparing the morpho-syntactic and phonetic performance of L2 learners, heritage bilinguals and monolingual speakers. The first study concentrated on different areas of the morpho-syntactic knowledge about clitic object pronouns and the second one focused on the global accent of Portuguese speech.

The results of both experiments reveal that HSs and L2ers of European Portuguese differ with respect to language properties that are acquired implicitly through extensive contact with the spoken language in early childhood, which is the case of pronunciation. Although heritage bilinguals showed higher variation in the global accent task than monolinguals, the majority was rated as native speakers, being clearly set apart from the L2ers, who

were consistently perceived as having a non-native accent. This is in line with studies indicating that early exposure to a language, even with the strong presence of a majority language, gives advantages in the acquisition of phonology compared to the acquisition of a language in post-puberty age (Au et al., 2002). On the one hand, this does not mean that attaining high proficiency in pronunciation is not possible and that perceptual and productive abilities to learn non-native segmental and suprasegmental contrasts become unavailable throughout the lifespan (Birdsong, 2007). Since the 1980s, several studies investigating the effects of perceptual training on both L2 speech perception and production have provided strong evidence of the plasticity of adult learners' mature perceptual systems by showing that phonetic training is effective in the modification of learners' perceptual patterns and in the improvement of their pronunciation accuracy (e.g., Iverson, Pinet & Evans, 2012; Rato, 2014). This suggests that the L2ers can develop a near-native-like accent with more extensive input (Piske, MacKay, & Flege, 2001) and phonetic training (Piske, 2007). On the other hand, early exposure to a language is not necessarily a sufficient condition for attaining native-like pronunciation (Flege, Frieda, & Nozawa, 1997). As other studies on HSs' phonological competence have documented (e.g. Godson, 2004; Kupisch et al., 2014; Stangen et al., 2015), early bilinguals may show an interaction of their phonemic systems, especially if the exposure to both languages is very unbalanced (with a much stronger exposure to the majority language). In our case, significant exposure to the HL in early years has contributed to the development of a native accent.

In the morpho-syntactic test, the HSs performed similarly to the monolinguals in experimental conditions testing clitic climbing and topicalization. The first structure is more frequent in the spoken language and the second structure shows an intrinsic case asymmetry that is part of the implicit knowledge of native speakers, provided by exposure to spoken input. Therefore, both properties are more challenging for L2ers, who have less exposure to naturalistic input in EP than HSs. These differences reveal that the linguistic competence of late L2ers and HSs relies on different input sources. On the other hand, heritage bilinguals have difficulties with structures that are acquired late in monolingual L1 acquisition, show variation, and are explicitly taught at school. Although the HSs included in this study had some degree of formal instruction in EP through HL courses attended until adolescence, it seems that this contact was not sufficient to fully stabilize their knowledge in this domain (i.e., with respect to allomorphic clitic forms). Nevertheless, there was a positive correlation between the HSs length of enrollment in HL courses and their overall performance, indicating that more contact with formal registers fosters HSs' knowledge of these properties. It remains open whether L2ers could acquire intrinsic asymmetries and variation typical of spoken EP under more intensive exposure to the spoken register. In order to disentangle the amount and type of input from the age factor (AoA), L2ers who acquire their second language exclusively under naturalistic conditions should be included in future research.

To conclude, this study showed that type of input and AoA are in fact crucial variables that shape linguistic performance in different ways. In line with claims made by several authors (Gujarro-Fuentes & Schmitz, 2015; Kupisch & Rothman, 2016; Pascual y Cabo & Rothman, 2012; Putnam & Sánchez, 2013; Rothman & Treffers-Daller, 2014), we think that these, and other variables, may make HS grammars different from monolingual grammars; however, these divergent grammars are not the result of acquisitions deficits. They are grammars in their own right.

REFERENCES

- Albirini, A., & Benmamoun, E. (2014). Aspects of second-language transfer in the oral production of Egyptian and Palestinian heritage speakers. *International Journal of Bilingualism*, 18(3), 244–273.
- Alarcón, I. (2011). Spanish grammatical gender under complete and incomplete acquisition: Early and late bilinguals' linguistic behavior within the noun phrase. *Bilingualism: Language and Cognition*, 14(3), 332-350.
- Andrade, A. (2010). The application of clitic climbing in European Portuguese and the role of register. In C. Borgonovo, M. Español-Echevarría, & P. Prévost (eds.), *Selected Proceedings of the 12th Hispanic Linguistics Symposium* (pp. 97-108). Somerville, MA: Cascadilla.
- Au, T., Knightly, L., Jun, S., & Oh, J. (2002). Overhearing a language during childhood. *Psychological Science*, 13(3), 238-243.
- Au, T., Knightly, L., Jun, S., Oh, J., & Romo, L. (2008). Salvaging a childhood language. *Journal of Memory and Language*, 58(4), 998–1011.
- Benmamoun E., Montrul S., & Polinsky, M. (2013). Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics*, 39(3/4), 129-181.
- Birdsong, D. (2007). Nativelike pronunciation among late learners of French as a second language. In O.-S. Bohn, & M. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 301-314). Amsterdam, Netherlands/Philadelphia, PA: John Benjamins.
- Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistic Analysis*, 20(1/2), 3–49.
- Bowles, M. (2011). Measuring implicit and explicit linguistic knowledge: What can heritage language learners contribute? *Studies in Second Language Acquisition*, 33(2) 247 – 272.
- Brito, A. M. (2008). Grammar variation in the expression of verb arguments: the case of Portuguese indirect object, *Phrasis. Studies in Language and Literature*, 49(2), 31-58.
- Cardinaletti, A., & Starke, M. (1999). The typology of structural deficiency. A case study of the three classes of pronouns. In H. van Riemsdijk (Ed.), *Clitics in the Languages of Europe* (pp. 145-233). Berlin: Mouton-De Gruyter.
- Chang, C. B., Yao, Y., Haynes, E. F., & Rhodes, R. (2011). Production of phonetic and phonological contrast by heritage speakers of Mandarin. *The Journal of the Acoustical Society of America*, 129(6), 3964-3980.
- Cuza, A., & Frank, J. (2015). On the role of experience and age-related effects: Evidence from the Spanish CP. *Second Language Research*, 30(1), 3-28.
- De Leeuw, E., Schmid, M., & Mennen, I. (2010). The effects of contact on native language pronunciation in an L2 migrant setting. *Bilingualism: Language and Cognition*, 13(1), 33-40.
- De Houwer, A. (1995). Bilingual language acquisition. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of child language* (pp. 219–250). Oxford, UK: Blackwell.
- DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(4), 499-534.

- Duarte, I. (1987). *Construção de Topicalização na Gramática do Português: Regência, Ligação e Condições sobre Movimento* (unpublished doctoral dissertation), Universidade de Lisboa, Portugal.
- Flege, J. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–273). Timonium, MD: York Press.
- Flege, J., & Fletcher, K. (1992). Talker and listener effects on the perception of degree of foreign accent. *Journal of the Acoustical Society of America*, 91(1), 370-389.
- Flege, J., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*, 25, 169-186.
- Flege, J., & MacKay, I. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, 26, 1-34.
- Flege, J., Munro, M., & MacKay, I. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97(5), 3125-3134.
- Flege, J., Yeni-Komshian, G., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, 41, 78–104.
- Flores, C. (2015). Understanding heritage language acquisition. Some contributions from the research on heritage speakers of European Portuguese. *Lingua. Special Issue in Fundamentally (in)complete acquisition*, 164(b), 251—265.
- Flores, C., Rinke, E., & Azevedo, C. (2017). Object realization across generations. A closer look on the spontaneous speech of Portuguese first and second generation migrants. In E. Domenico (Ed.), *Complexity in acquisition* (pp. 179-206). Newcastle upon Tyne, UK: Cambridge Scholars Press.
- Fowler, C., Sramko, V., Ostry, D., Rowland, S., & Hallé, P. (2008). Cross-language phonetic influences on the speech of French-English bilinguals. *Journal of Phonetics*, 36(4), 649-663.
- Godson, L. (2004). Vowel Production in the Speech of Western Armenian heritage speakers. *Heritage Language Journal*, 2, 44-69. Available from <http://heritagelanguages.org>
- Goldstein, B. (2011). *Cognitive Psychology: Connecting mind, research, and everyday experience* (3rd ed.). Belmont, CA: Wadsworth.
- Guijarro-Fuentes, P., & Schmitz, K. (2015). The nature and nurture of heritage language acquisition [editorial]. *Lingua*, 164(B), 239–250.
- Guion, S. (2003). The vowel systems of Quichua-Spanish bilinguals. *Phonetica*, 60, 98–128.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26, 55–88.
- Hopp, H., & Schmid, M. (2013). Perceived foreign accent in L1 attrition and L2 acquisition: the impact of age of acquisition and bilingualism. *Applied Psycholinguistics*, 34(2), 361-394.
- Hyltenstam, K., & Abrahamsson, N. (2003). Maturational constraints in SLA. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 539-588). Malden, MA: Blackwell.
- Hyltenstam, K., Bylund, E., Abrahamsson, N., & Park, H-S. (2009). Dominant-language replacement: The case of international adoptees. *Bilingualism: Language and Cognition*, 12(2), 121-140.

- Iverson, P., Pinet, M., & Evans, B. (2012). Auditory training for experienced and inexperienced second language learners: Native French speakers learning English vowels. *Applied Psycholinguistics*, 33(1), 145-60.
- Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. London, UK: Routledge.
- Jesney, K. (2004). *The use of global foreign accent rating in studies of L2 acquisition*. Calgary, Alberta: University of Calgary Language Research Centre Reports.
- Keating, G., VanPatten, B., & Jegerski, J. (2011). Who was walking on the beach? Anaphora resolution in Spanish heritage speakers and adult second language learners. *Studies in Second Language Acquisition*, 33(2), 193– 221.
- Knightly, L., Jun, S., Oh, J., & Au, T. (2003). Production benefits of childhood overhearing. *Journal of the Acoustic Society of America*, 114, 465-474.
- Kupisch, T. (2012). Specific and generic subjects in the Italian of German–Italian simultaneous bilinguals and L2 learners. *Bilingualism: Language and Cognition*, 15(4), 736-756.
- Kupisch, T. & Rothman, J. (2016). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*. OnlineFirst. doi: 10.1177/1367006916654355
- Kupisch, T., Barton, D., Klaschik, E., Lein, T., Stangen, I., & Weijer, J. (2014). Foreign accent in adult simultaneous bilinguals. *Heritage Language Journal*, 11(2), 123-150. Available from <http://heritagelanguages.org>
- Lleó, C. (2006). The acquisition of prosodic word structures in Spanish by monolingual and Spanish-German bilingual children. *Language and Speech*, 49(2), 205-229.
- Long, M. (1990). Maturation constraints on language development. *Studies in Second Language Acquisition*, 12(3), 251-285.
- Mack, M. (1989). Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals. *Perception & Psychophysics*, 46(2), 187–200.
- MacKay, I. R. A., Flege, J. E., Piske, T., & Schirru, C. (2001). Category restructuring during second-language speech acquisition. *Journal of the Acoustical Society of America*, 110(1), 516-528.
- Meisel, J. M. (2011). *First and second language acquisition: Parallels and differences*. Cambridge, UK: Cambridge University Press.
- Montrul, S. (2010). Current issues in heritage language acquisition. *Annual Review of Applied Linguistics*, 30, 3-23.
- Montrul, S. (2011). Morphological errors in Spanish second language learners and heritage speakers. *Studies in Second Language Acquisition*, 33(2), 163–192. doi:10.1017/S0272263110000720
- Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge, UK: Cambridge University Press.
- Montrul, S., Foote, R., & Perpiñán, S. (2008): Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning* 58(3), 503–553.
- Montrul, S., Davidson, J., de la Fuente, I., & Foote, R. (2014). Early language experience facilitates the processing of gender agreement in Spanish heritage speakers. *Bilingualism. Language and Cognition*, 17(1), 118-138.
- Moyer, A. (2004). *Age, accent and experience in second language acquisition. An integrated approach to critical period inquiry*. Clevedon, UK: Multilingual Matters.
- Munro, Murray J., & Derwing, T. M. (1994). Evaluations of foreign accent in extemporaneous and read material. *Language Testing*, 11(3), 253-266.

- O'Grady, W., Lee, M., & Choo, M. (2001). The acquisition of relative clauses by heritage and non-heritage learners of Korean as a second language: A comparative study. *Journal of Korean Language Education*, 12, 283-94.
- Oh, J., Jun, S.-A., Knightly, L., & Au, T. (2003). Holding on to childhood language memory. *Cognition*, 86(B), 53–B64.
- Paradis, J. (2001). Do bilingual two-year-olds have separate phonological systems?. *International Journal of Bilingualism*, 5(1), 19 – 38.
- Pascual y Cabo, D., & Rothman, J. (2012). The (il)logical problem of heritage speaker bilingualism and incomplete acquisition. *Applied Linguistics*, 33(4), 450 - 455.
- Pires, A., & Rothman, J. (2009). Disentangling sources of incomplete acquisition: An Explanation for competence divergence across heritage grammars. *International Journal of Bilingualism*, 13(2), 211-39.
- Piske, T. (2007). Implications of James E. Flege's research for the foreign language classroom. In O.-S. Bohn, & M. Munro (Eds.), *Language experience in second language speech learning. In honor of James Emil Flege* (pp. 301-14). Amsterdam, Netherlands/Philadelphia, PA: John Benjamins.
- Piske, T., Flege J. E., MacKay, I. R. A., & Meador, D. (2002). The production of English vowels by fluent early and late Italian-English bilinguals. *Phonetica*, 59, 49 - 71.
- Piske, T., MacKay, I., & Flege, J. (2001). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, 29, 191-215. Retrieved from http://jimflege.com/files/Piske_MacKay_factors_affecting_JP_2001.pdf
- Putnam, M., & Sánchez, L. (2013). What's so incomplete about incomplete acquisition? A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism*, 3(4), 476-506.
- Rato, A. (2014). Effects of perceptual training on the identification of English vowels by native speakers of European Portuguese. *Proceedings of the International Symposium on the Acquisition of Second Language Speech. Concordia Working Papers in Applied Linguistics*, 5, 529 – 546.
- Rauber, A., Rato, A., Kluge, D., & Santos, G. (2012). *TP-S (Version 3.1)*. [Software]. Pelotas, Brazil: Worken.
- Rinke, E. & Flores, C. (2014). Heritage Portuguese bilinguals' morphosyntactic knowledge of clitics. *Bilingualism. Language and Cognition*, 17(4), 681 - 699. doi:10.1017/S136672891300076X
- Rothman, J. & Treffers-Daller, J. (2014). A Prolegomenon to the construct of the native speaker: Heritage speaker bilinguals are natives too! *Applied Linguistics*, 35(1), 93-98.
- Santos, A. L., & Flores, C. (2016). Comparing heritage speakers and late L2-learners of European Portuguese: verb movement, VP ellipsis and adverb placement. *Linguistic Approaches to Bilingualism*, 6(3), 308 - 340.
- Stangen, I., Kupisch, T., Ergün Proietti, A. L., & Zielke, M. (2015). Foreign accent in heritage speakers of Turkish in Germany. In H. Peukert (Ed.), *Transfer effects in multilingual language development*, pp. 87–108. Amsterdam, Netherlands: John Benjamins.

NOTES

- [1.](#) A Bonferroni correction was applied to all Mann-Whitney tests throughout the paper; therefore, the statistical results for this test will be reported at a .017 level of significance.
- [2.](#) *Perfil Sociolinguístico da Fala Bracarense*, (coord. Pilar Barbosa), Universidade do Minho, Braga.