Running-head: Language dominance switch in Galician bilinguals

Title: The effects of language dominance switch in bilinguals: Galician new speakers' speech production and perception

Authors: Gisela Tomé Lourido[1,2] & Bronwen G. Evans[2]

Affiliation: [1]University of Leeds; [2]Speech, Hearing & Phonetic Sciences, University College London

Acknowledgements:

We thank Antonio Cardenal for his help in providing and implementing the forced aligner. We are grateful to Xosé Luís Regueira and Elisa Fernández Rei for help with data collection and discussions about this project. The writing of this article also benefitted from networking with members of COST Action IS1306 New Speakers in a Multilingual Europe: Opportunities and Challenges and discussions at the Sociophonetics Discussant session at ICPhS 2015 led by Paul Foulkes.

Address for correspondence:

g.tomelourido@leeds.ac.uk

School of Languages, Cultures and Societies

Room B14, Michael Sadler Building

University of Leeds

Leeds LS2 9JT

United Kingdom

Keywords: language dominance switch, speech production, speech perception, new speakers

Abstract:

It has long been debated whether speech production and perception remain flexible in adulthood. The current study investigates the effects of language dominance switch in Galician new speakers (*neofalantes*) who are raised with Spanish as a primary language and learn Galician at an early age in a bilingual environment, but in adolescence, decide to switch to using Galician almost exclusively, for ideological reasons. Results showed that *neofalantes* pattern with Spanish-dominants in their perception and production of mid-vowel and fricative contrasts, but with Galician-dominants in their realisation of unstressed word-final vowels, a highly salient feature of Galician. These results are taken to suggest that despite early exposure to Galician, high motivation and almost exclusive Galician language use post-switch, there are limitations to what *neofalantes* can learn in both production and perception, but that the hybrid categories they appear to develop may function as opportunities to mark identity within a particular community.

1. Introduction

Learning the sounds of a new language in adulthood is often very difficult. The sound system of the first language (L1) we learn influences the acquisition of subsequent languages (Flege, 1995; Best, 1994). This is the case even in bilingual contexts, where individuals have early exposure to their second language (L2; e.g., from early childhood) and listen to it on an everyday basis. Such bilinguals, dominant in one language, often find it very difficult to acquire phonetic categories that do not exist in their second, non-dominant language (Pallier, Bosch & Sebastián-Gallés, 1997; Sebastián-Gallés & Soto-Faraco, 1999). Some have accounted for this difficulty by arguing for a lack of behavioural plasticity (e.g., Pallier et al., 1997). However, others have argued that difficulties in acquiring new phonetic categories in an L2 result not from a loss of plasticity but from a lack of use of the L2. For example, Flege & Mackay (2004) investigated the perception of English vowels by native speakers of Italian. They found that early learners who reported using their L1 (i.e., Italian) seldom resembled native English speakers in terms of their vowel discrimination, whilst those who used their L1 often, did not, suggesting that continued usage of the L1 affected acquisition of the L2. Indeed, theories of L2 perception, such as the Perceptual Assimilation Model (PAM: Best, 1994, 1995; PAM-L2: Best & Tyler, 2007) and the Speech Learning Model (SLM; Flege, 1995), have proposed that L2 sounds are filtered through L1 phonetic categories. According to these theories, failing to create new phonetic categories is not a consequence of a reduction in neural plasticity. Rather, the mechanisms used for learning remain intact throughout the lifespan (Flege, 1995) and perceptual learning continues into adulthood (Best, 1995), but continued use of the L1 means that changes in perceptual processing due to language experience are reinforced,

making it harder to acquire the L2 (Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Kettermann and Siebert, 2003).

The current study aims to further investigate whether speech production and perception remain flexible across the lifespan by focussing on a different bilingual population. Similar to the participants in Pallier et al. (1997), this group grew up in a bilingual environment, where they were exposed to both Galician and Spanish from an early age on a daily basis, but were dominant in one language, Spanish. In adolescence, though, they decided to switch to using their non-dominant language, Galician, predominantly or exclusively, for ideological reasons. This group are known as *neofalantes* (new speakers). *Neofalantes*, who have early and extensive exposure to the non-dominant language, almost exclusive use of their new language post-switch and are highly motivated, constitute an ideal population to examine whether experience with the L2, together with early and extensive exposure leads to the formation of new, native-like phonetic categories.

1.1. Neofalantes as unbalanced bilinguals

Previous research has shown that even simultaneous bilinguals who were exposed to both languages before the age of 1 year, and pass as native speakers of both languages do not perceive speech like monolingual speakers in one of their languages (Cutler, Mehler, Norris & Segui, 1989, 1992; Sebastián-Gallés, Echeverría & Bosch, 2005; Dupoux, Peperkamp & Sebastián-Gallés, 2010), suggesting that there is always a dominant language for optimal processing. Likewise, early bilinguals may also be dominant in one of their languages. Performance with the non-dominant language is task-dependent: early and late bilinguals tend to perform like native listeners on tasks that involve pre-lexical processing (e.g., categorisation or phoneme identification), but not in tasks that tap into lexical processing (Sebastián-Gallés & Díaz, 2012). This

preference for the dominant language has been shown for different aspects of language processing, e.g., cue sensitivity (Hazan & Boulakia, 1993) and speech segmentation (Cutler et al., 1989).

This is the case even when bilinguals live in a context where they have continuous exposure to both languages, e.g., Catalonia. For example, Pallier et al. (1997) tested highly proficient early bilinguals who had learned either Spanish or Catalan from birth. Catalan has a phonemic contrast between the mid-vowels /ɛ/-/e/, while Spanish has only one front mid-vowel /e/. Results from identification and discrimination tasks showed that participants who had learnt Catalan from birth had two distinct phonemic categories. Additionally, when participants were asked to rate the typicality of different vowels in two Catalan and one Spanish word, Catalan-dominant participants produced the expected responses for the Catalan vowels but conflated Spanish and Catalan /e/ for the Spanish word. However, Spanish-dominant listeners behaved differently; though they had some awareness of the existence of the two different Catalan vowels, they gave different typicality judgements from Catalan-dominant bilinguals. Further research (e.g., Sebastián-Gallés & Soto-Faraco, 1999; Bosch, Costa & Sebastián-Gallés, 2000) has shown that Spanish-dominant bilinguals perform more poorly than Catalan-dominants in perception tasks with contrasts that only exist in Catalan and not in Spanish (e.g., front and back mid-vowels /ɛ/-/e/, /ɔ/-/o/ and fricative contrasts /s/-/z/, /ʃ/-/ʒ/), and that lack of sensitivity to the non-dominant language contrast extends to lexical representations; in a lexical decision task Spanish-dominants performed as well as Catalan-dominants but they processed some Catalan minimal pairs as homophones (Pallier, Colomé & Sebastián-Gallés, 2001). These differences become evident in childhood (Ramon-Casas, Swingley, Sebastián-Gallés & Bosch, 2009), such that acquiring phonetic contrasts in the non-dominant language appears extremely difficult,

and the malleability of L1 phonetic categories severely limited, even with early and extensive exposure to the language (Sebastián-Gallés & Soto-Faraco, 1999).

Work by Flege and colleagues on sequential bilinguals who moved countries and became dominant in their L2 challenged this view. Flege, MacKay and Meador (1999) compared different groups of Italian learners of English and found that early learners who moved to Canada around 7 years old, did not differ significantly from English native monolinguals in their discrimination of English vowels. They interpreted this as indicating that early bilinguals were able to create new phonetic categories in their L2 (see also Flege & Mackay, 2004 for Italian learners of English, and Mora, Keidel & Flege, 2011; 2015 for Catalan-Spanish bilinguals). This indicates that attunement to the L1 does not prevent early bilinguals from performing like native monolinguals in terms of perception of their L2, and instead suggests that variation in accuracy in perception may be largely determined by patterns of L2 language use.

Research on speech production has likewise shown conflicting results. Some studies support the idea that dominant bilinguals have language-specific phonetic categories in production; however, they may not have monolingual-like realisations in their non-dominant language (e.g., Hazan & Boulakia, 1993). For example, Amengual (2014, 2016) showed that Spanish-dominant bilinguals in Majorca were able to maintain a contrast between the Catalan front /ɛ/-/e/ and back /ɔ/-/o/ mid-vowels, although the contrasts were smaller than those of Catalan-dominant bilinguals. Similarly, in production of the alveolar lateral approximant /l/, which differs in the degree of velarisation in Spanish and Catalan, both Spanish- and Catalan-dominant bilinguals had language-specific realisations, which were different from those used by the dominant group who had learned the language from birth (Simonet, 2010). That is, production of the Catalan and Spanish variants differed according to the language being tested (i.e.,

Catalan, Spanish), and whether this was the speaker's dominant language. However, Simonet (2011) found that Spanish-dominant speakers in Majorca (i.e., Spanish-Catalan bilinguals) had a merged contrast for /ɔ/-/o/. Similarly, for Galician, Spanish-dominant speakers have difficulty maintaining a phonetic contrast between the front and back mid-vowels (Amengual & Chamorro, 2015). Like Catalan, Galician contrasts open-mid and close-mid front /ɛ/-/e/ and back /ɔ/-/o/ vowels in stressed position. Amengual and Chamorro (2015) tested Galician-dominant and Spanish-dominant speakers in their perception and production of the front and back mid-vowels. The results showed that Spanish-dominant listeners had greater difficulty identifying and discriminating the mid-vowels, while Galician-dominants had robust categorical perception of the contrasts, as expected. In production, Spanish-dominants also had difficulty maintaining this contrast and had a merged category for front vowels, though they appeared to have a small contrast for back vowels. Conversely, Galician-dominant speakers had a distinct contrast for both front and back vowel pairs.

1.2. Language learning in its social context

One factor that has not received much attention in speech learning research is the role of social factors. Language use in minority language communities, such as Galicia, is likely further complicated by the influence of the speaker's attitudes towards the languages they choose to use. *Neofalantes* often switch dominance to the minority language for ideological reasons; thus, it is possible that their speech production might not only be accounted for in terms of language learning constraints, but may also be influenced by social factors, e.g., identity. Research in sociophonetics has shown that these social factors affect the use of phonetic variables within one language (Eckert, 2000; 2008; Foulkes & Docherty, 2006) and that speakers may alter the phonetic features they use to show belonging or identification with a particular group (Evans &

Iverson, 2004; 2007). For example, Evans and Iverson (2007) examined speech perception and production in a group of students who moved from a small community in the Midlands (in the centre of England) to study at university, where they encountered speakers of different accents, in particular the standard variety. Although these students retained certain phonetic variants, e.g., to show belonging to their home community, they changed their production of others to better fit their new multidialectal community. Some, but not all, changed their production of the STRUT vowel /ʌ/ (which in their native accent is merged with the FOOT vowel /ʊ/) to make it more centralised. However, their realisation of this phoneme was not the same as that of native speakers of the standard accent, nor were these shifts accompanied by changes in perception. This suggests that whilst speakers might be able to change certain aspects of their speech production at a relatively late stage in their language development, late adolescence, there are limits to this flexibility.

Less is known about how bilinguals encode identity through the use of their languages. Recent work with L2 learners has highlighted the importance of identity in acquiring regional features in an L2; Polish migrants in Manchester were less likely to produce the local variant of (ing) if they were planning on returning to Poland, and more likely to adopt the local variant if they were planning on remaining in Manchester (Drummond, 2012a). Likewise, the more positive the speaker's attitude towards Manchester, the more likely they were to merge the STRUT and FOOT vowels, producing both with /ʊ/ (i.e., the local variant) rather than with the standard Southern British English /ʌ/-/ʊ/ split that they had typically been taught in school (Drummond, 2012b). Given that those who planned to stay in Manchester were more likely to identify positively with the city, one possibility is that they used these variables to signal belonging to their host community.

For *neofalantes*, the switch in language dominance is not a result of external factors, e.g., moving countries, and as such, cannot be disentangled from or understood without reference to the community to which they belong. New speakers have been documented in other minority language communities in Europe (cf. O'Rourke, Pujolar & Ramallo, 2015), though the definition of new speaker covers very different types of individuals; new speakers range from low or high proficiency L2 adult learners (Nance, McLeod, O'Rourke & Dunmore, 2016) to bilinguals in immersion schooling (Nance, 2015). Most Galician new speakers are bilinguals who learn Spanish as their home language and Galician outside the home. This is possible because most of the Galician population is bilingual and has a high competence in both languages. Note that high competence in Galician does not necessarily come from schooling; it may also come from acquiring the language from the environment (Ramallo & O'Rourke, 2014), e.g., through grandparents, friends or the wider community. O'Rourke and Ramallo (2015, p. 148) define *neofalantes* as "individuals for whom Spanish was their language of primary socialisation, but who at some stage in their lives (usually early to late-adolescence) have adopted Galician language practices and on occasions displaced Spanish all together". Additionally, this language displacement is often motivated by ideological, political or socio-cultural factors. Indeed, Ramallo (2013) situates the origin of this group of speakers in the 1980s, after Spain's transition to democracy. This transition had far-reaching implications for the sociolinguistic situation in Galicia, including the Galician language gaining co-official status, its recognition as a symbol of Galician identity, and the start of the standardisation process, through which a standard Galician language variety was developed (Ramallo, 2013). This standard variety was then incorporated into spaces that were exclusive to Spanish, e.g., education, the media and public administration, which facilitated access to Galician for non-traditional speakers

(i.e., those who had not learned Galician from birth). Nevertheless, although *neofalantes* are often characterised as being urban middle class speakers (O'Rourke & Ramallo, 2013b), they are also found in rural environments and across different social classes. Investigating the effect of origin, i.e., urban or rural, is not a central aim of this study. However, this factor might influence the accent of Galician acquired; bilinguals growing up in urban areas, characterised by a higher usage of Spanish (IGE, 2013), are likely to be exposed more often to Spanish and Spanish-accented varieties of Galician than those in rural areas.

Although several classifications for *neofalantes* have been proposed, in this article we define a *neofalante* according to the following three characteristics:

1) Early experience with the minority language: although speakers only used Spanish with their parent(s) and vice versa, they learned Galician as children, either through school, friends, the extended family or the wider community.
2) There is a long-term switch in language dominance: speakers changed from being dominant in Spanish to displace this language either predominantly or totally to speak Galician (almost) exclusively.
3) Motivations for language switch: this switch takes place due to ideological, political or socio-cultural motivations. These speakers are normally committed to the revitalisation of the Galician language.

This definition is more restrictive than those used for new speakers in other European minority communities (e.g., Scotland: Nance, 2015, Nance et al., 2016; Corsica: Jaffe, 2015; Catalonia: Pujolar & Puigdevall, 2015), but matches the majority use of this label in the Galician case (Ramallo, 2013; Ramallo & O'Rourke, 2014). Although

neofalantes thus represent a small proportion of the total population[1], O' Rourke and Ramallo (2015) have argued for "*neofalantismo*" as a social movement, with *neofalantes* an active minority. An active minority is one in which "individuals or groups [...] through their behaviour attempt to influence both the attitudes and practices of the majority and in doing so, bring about social change" (p. 151). Consequently, these authors suggest that becoming a new speaker "requires innovative action through an appropriation of a new linguistic space as well as commitment to the transformation of society from below" (p. 153).

1.3. The current study

The current study investigates the consequences of a long-term language switch for speech production and perception in *neofalantes*, Spanish-dominant bilinguals in Galicia who consciously switch in adolescence from using Spanish to Galician, predominantly or exclusively, for ideological reasons. We investigate *neofalantes*' production and perception of two mid-vowel contrasts in stressed position /ɛ/-/e/ and /ɔ/-/o/ and a sibilant consonant contrast /s/-/ʃ/, all of which do not exist in Spanish, as compared to Spanish-dominant and Galician-dominant bilinguals. We also test the production of reduced word-final vowels, which have a distinctive phonetic realisation in Galician and have been reported to contribute to the perception of the Galician accent (Regueira, 2012).

As previously mentioned, Galician has a phonemic contrast between mid-front and back vowels which does not exist in Spanish, and so based on previous research, we

[1] In 2008, 24,216 people whose initial language was Spanish switched to speak only Galician or more Galician than Spanish by personal decision (IGE, 2008). If we understand that this figure represents *neofalantes*, they would form around 2% of the Spanish-dominant population (1,105,486) and a little less of the Galician-dominant population (1,466,915 people).

predict these pairs of vowels will be difficult for *neofalantes* to perceive and produce. The fricatives /s/-/ʃ/ are different phonemes in Galician, but only /s/ exists in Spanish. There are no descriptions of the production and perception of this contrast by either Spanish-dominant bilinguals or *neofalantes*, but descriptions of urban varieties often associated with these groups of speakers report apical realisations of /ʃ/ (Regueira, 1999; González González, 2008), suggesting that for these two groups /ʃ/ may be more similar to /s/. Based on such impressionistic accounts, the fricative contrast may be difficult for *neofalantes* to perceive and produce. Finally, unstressed word-final vowels are raised and centralised in Galician (Molinos Castro 2002; Regueira, 2007), but not in standard Spanish. It has been reported that word-final vowels are also reduced in the variety of Spanish spoken in Galicia (Rojo, 2005), but others have claimed that Spanish-dominant speakers do not produce reduced vowels. For example, the accent used in the media, associated with the variety spoken by Spanish-dominant speakers, has been claimed to have unreduced word-final vowels, with a similar intensity and duration to that of stressed vowels (Regueira, 1994). Assuming that Galician-dominant and Spanish-dominant speakers behave differently from each other in their production of this feature, we investigate where *neofalantes* lie on this continuum.

2. Experiment 1: Measurement of Production

2.1. ***Method***

Participants

Sixty-eight participants were tested. Participants were recruited from the University of Santiago de Compostela, which has the largest and most heterogeneous student population in Galicia. This facilitated recruitment of participants from different backgrounds, i.e., *neofalantes*, Galician-dominant, Spanish-dominant. Three

participants were excluded because they did not meet the criteria for the experiment. The remaining 65 participants grew up in Galicia, had not lived anywhere else for more than a year and were bilingual in Galician and Spanish.

Participants were all students and at the time of recruitment were 18–30 years old (median 20 years). After the experiment, they completed a detailed language background questionnaire which included questions about language background and exposure, language use, and social variables (see Appendix A). This was used to classify participants into three groups, resulting in 14 *neofalantes* (7 female), 22 Galician-dominant (12 female), 20 Spanish-dominant (12 female) and 6 simultaneous bilinguals (3 female). The data from the simultaneous bilinguals, who were raised in a one-parent one-language setting, will not be presented here. The criteria used to assign participants to the different groups was as follows:

- *Neofalantes*: raised predominantly in Spanish (i.e., their parent(s) used to speak to them in Spanish), but decided to adopt Galician as their dominant language in adolescence (13-20 years old, median 17) for ideological or cultural reasons. Since this switch, they have mainly spoken Galician (mean reported Galician use = 4.65/5)[2].
- Galician-dominant bilinguals: raised predominantly in Galician (i.e., their parent(s) spoke Galician to them) and have always spoken mainly Galician (mean reported Galician use = 4.64/5).

[2] Question 31 in the Language Background Questionnaire (Appendix A).

- Spanish-dominant bilinguals: raised predominantly in Spanish (i.e., their parent(s) spoke Spanish to them) and have always spoken mainly Spanish (mean reported Galician use = 2.36/5).

A further 3 participants who did not meet any of these criteria were also excluded, giving a final total of 56. Note that all participants were raised in a bilingual community and, thus, would have been exposed to both languages from an early age. For 51 participants, both parents had been born and raised in Galicia and in 5 cases one of the parents had been born in Spain (1 *neofalante*, 2 Galician-dominants, 2 Spanish-dominants), but all participants had at least one parent who had been born in Galicia. Therefore, the only difference between Spanish-dominants and *neofalantes* in terms of language background was that *neofalantes* made a conscious decision in adolescence to always speak Galician. The smaller sample size in the *neofalantes* group is due to various constraints related to recruitment. First, as mentioned in the Introduction, *neofalantes* constitute a small proportion of the Galician population: less than 2%. Additionally, the label '*neofalante*', also a folk term used in the community, may have negative connotations in certain contexts and *neofalantes* themselves may or may not identify with it (see O'Rourke & Ramallo, 2013a, 2015, for a detailed description of *neofalantes*' sociolinguistic profile), making its use inappropriate for recruiting purposes. Finally, recruiting participants by enquiring about their language background would direct their attention to the Galician language, and might have made *neofalantes* feel they were being assessed. Therefore, participants were only asked general questions before the experiment, and *neofalantes* were recruited by sampling the population or targeting certain groups.

Participants came from both urban and rural backgrounds (*neofalantes*: 8 urban, 6 rural; Galician-dominant: 5 urban, 17 rural; Spanish-dominant: 11 urban, 9 rural). The

imbalance in the Galician-dominant group does not permit a reliable interpretation of the effect of origin, but the results could form the basis of future research. None of the subjects reported any speech, hearing or language disorders at the time of testing.

Materials

The stimuli consisted of a wordlist and a text that contained all three variables of interest; mid-vowels, sibilant fricatives and word-final vowels. The subset of Galician words used for the mid-vowel analysis was *pazo* [ˈpaθo̯] 'pazo[3]', *peza* [ˈpɛθa̯] 'piece', *peto* [ˈpeto̯] 'pocket', *pita* [ˈpita̯] 'hen', *pote* [ˈpɔte̯] 'pot', *pozo* [ˈpoθo̯] 'well', *pucho* [ˈput͡ʃo̯] 'calf', *seca* [ˈseka̯] 'dry', *sota* [ˈsɔta̯] 'knave', *sopa* [ˈsopa̯] 'soup'. The target was the first, stressed vowel. For the fricative analysis, the words were *pase* [ˈpase̯] 'pass' and *paxe* [ˈpaʃe̯] 'page'. In this case, the target sound was the fricative. The analysis for the unstressed word-final vowels included all the words for the two previous analyses, as well as *pata* [ˈpata̯] 'paw', *sapo* ['sapo̯] 'toad', *saco* ['sako̯] 'sack bag', *sito* ['sito̯] 'situated', *suco* ['suko̯] 'furrow' and *pare* ['pare̯] 'stop'. The target was the final, unstressed vowel. Each of these words was recorded in phrase-final position in the carrier sentence *digo a palabra* ____ (I say the word ____) and in phrase-medial position in the carrier sentence *digo a palabra* ____ *con coidado* (I say the word ____ carefully). As Galician is closely related to Spanish, most of the stimuli used are cognates. The text was a modified version of "The North Wind and the Sun" (*O vento do norte e o sol*). The original text contained only a small number of key variables, and so a sentence was added to increase the number of instances of these, giving 3-6 repetitions of each target variable (see Appendix B). All recordings were made in Praat

[3] a type of Galician traditional house, similar to a manor house.

(Boersma & Weenink, 2014), in a quiet room using a Samson C01U microphone connected to a laptop, and with a sampling rate of 44.1kHz, 16-bit resolution.

Procedure

Participants recorded one repetition of the wordlist and the text. To equalize any accommodation effects across participants, all testing was carried out by the first author, who was also a bilingual speaker from Galicia, and the session was conducted in Galician. None of the participants had a close relationship with the experimenter, though they knew that she was from Galicia; all were university students or friends of friends who were naive to the goal of the experiments.

Recordings from both the wordlist and the text were segmented using a forced aligner for Galician (García-Mateo, Cardenal, Regueira Fernández, Fernández Rei, Martínez, Seara, Varela & Basanta Llanes, 2014) and any errors hand-corrected.

Three different sets of measurements were made for each of the three variables; mid-vowels, fricatives and word-final vowels. For the mid-vowel analysis, the mean F1 and F2 values were extracted using Praat scripts (Boersma & Weenink, 2014) from each target word. Measurements were taken from the 50% middle portion of the stressed vowel where formant values are most stable (average duration 85ms). Praat's default formant tracking settings were used. Only the mid-vowels /e ɛ o ɔ/ were included in the statistical analysis. The vowels /a i u/ were used in the normalisation procedure and are included in plots for reference. This gave 2-4 measurements per mid-vowel (i.e., /e ɛ o ɔ/) for the wordlist and 3-6 measurements per mid-vowel for the text. Formant measures that were 2 standard deviations outside the F1 or F2 mean per vowel were checked and hand-corrected if necessary. To be able to compare data from male and female talkers, measurements were normalised using the Lobanov method which has been shown to

reduce the effects of anatomical and physiological variation, whilst retaining phonemic variation (Adank, Smits & van Hout, 2004).

For the sibilant fricative analysis, the centre of gravity (CoG) was calculated in the middle portion (40ms around the midpoint) of the fricative (average duration 98ms) in each target word, using Praat (Boersma & Weenink, 2014). This gave 2 measurements per consonant (i.e., /s ʃ/) for the wordlist and 5-9 measurements per consonant for the text. Although other acoustic variables, such as skew and kurtosis, could contribute to differences in fricative production, CoG was chosen because it has been shown to differentiate place of articulation in fricatives, in particular for Galician (Regueira & Ginzo, in press); alveolar sibilants have been shown to have a higher spectral mean than post-alveolar sibilants (Jongman, Wayland & Wong, 2000; Regueira & Ginzo, in press). Fricatives produced by women have been shown to have higher spectral means than those produced by men (Jongman et al., 2000), perhaps due to biological differences. Galician sibilants are prototypically voiceless; however, given that the voiced counterparts do not occur in the language contrastively, some speakers may produce voiced sounds in certain contexts. Given that voicing may affect spectral moments (Jongman et al., 2000), segments which had a voiced portion longer than 20% of the total length of the sound were manually checked, and fricatives that were mostly or fully voiced were excluded from further analysis (16 tokens). Fricatives shorter than 40ms were also excluded (21 tokens). CoG measures that were 2 standard deviations outside the mean for each phoneme were checked and hand corrected if necessary (N=944; 485 alveolar, 459 post-alveolar phonemes).

Finally, word-final vowels were analysed in a similar way to mid-vowels. The mean F1 and F2 values were extracted using Praat scripts (Boersma & Weenink, 2014) from the 50% middle portion of the unstressed word-final vowel in each target word (average

duration 65ms). The same formant tracking settings were used. Only the mid unstressed vowels were included in the analysis; [ạ] was included for reference. This gave 8-16 measurements per vowel [ẹ ọ] for the wordlist and 10 measurements per vowel for the text. Formant measures that were 2 standard deviations outside the F1 or F2 mean per vowel were checked and hand corrected if necessary (N=1741). As before, data was then normalised using the Lobanov method (Adank et al., 2004).

2.2. ***Results***

Mid-vowels in stressed position

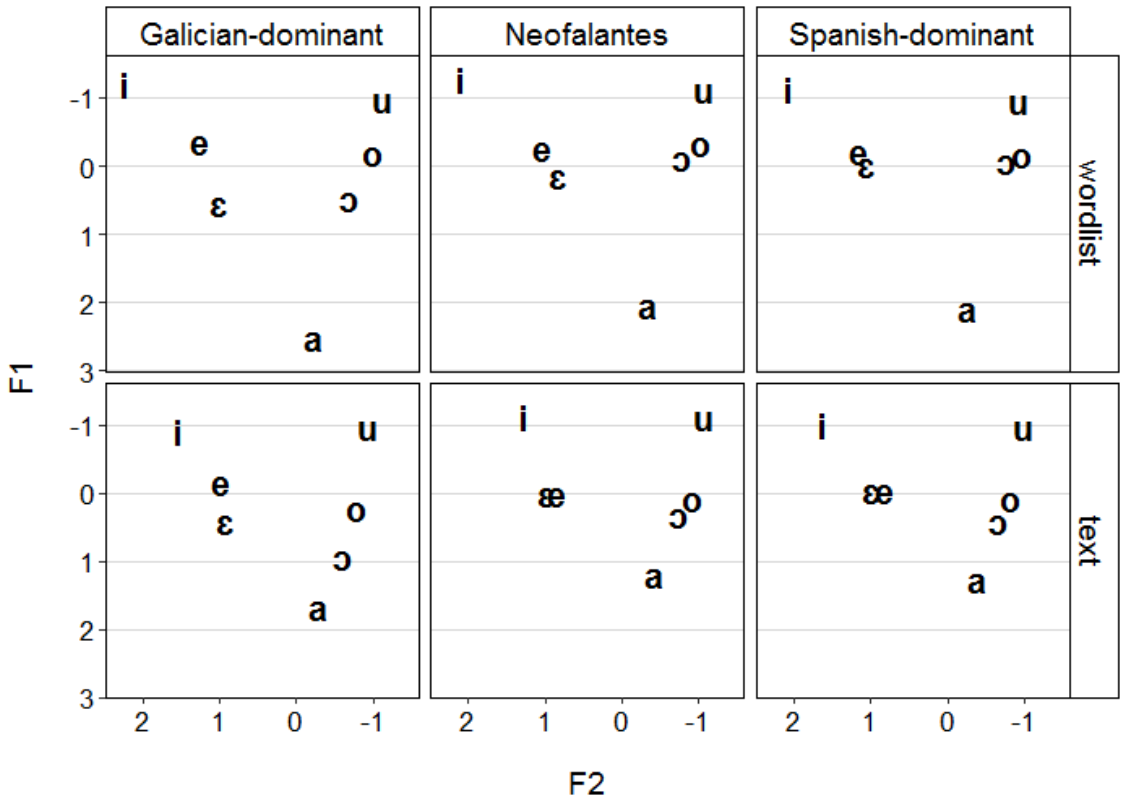

Figure 1: Average F1 and F2 formant frequencies by group (Galician-dominant,

neofalantes, Spanish-dominant) and speech style (wordlist, text). Plots display normalised values.

Figure 1 displays the mean normalised formant values (F1, F2) of test words in the wordlist and text as produced by the three groups, Galician-dominant, Spanish-dominant and *neofalantes*. To investigate whether the groups had a split category, the Euclidean distance was calculated separately for front and back vowels for each speaker for each speech style (wordlist, text; N=224). These values were used in all further mid-vowel analyses. Given that the Euclidean distance yields a skewed distribution, a rank-transformation to normality was applied to fit a linear mixed-effects regression model. The variables of group (*neofalantes*, Galician-dominant, Spanish-dominant speakers), style (wordlist, text), origin (rural, urban), vowel (front, back) and sex (female, male) were included as fixed factors in the model up to three-level interactions. All possible by-participant random slopes were included in the model, following Barr, Levy, Scheepers and Tily (2013). A simple coding scheme was used for this model and all others in this article, in which each level is compared to the reference level (similar to the treatment coding scheme), but the intercept represents the grand mean rather than the cell mean of the reference level. "*Neofalantes*" was selected as the reference level in the group factor to investigate whether this group behaved differently from the Galician-dominant and Spanish-dominant bilinguals. The p-values and degrees of freedom for this model and all the models in this article were estimated using the Satterthwaite approximation from the lmerTest package (Kuznetsova, Bruun Brockhoff & Haubo Bojesen Christensen, 2016) in R (R Core Team, 2016).

	Mid vowels Euclidean Distance	Sibilant fricatives CoG	Word-final vowels f1
Intercept	-2.081(46)*	n.s.	-1.930(34) .
MAIN EFFECTS			
Group (GD)	**2.996 (46) ****	n.s.	**n.s.**
Group (SD)	**n.s.**	n.s.	**2.061 (47) ***
Phoneme	n.s.	n.s.	n.s.
Style	3.453 (56) **	n.s.	-2.099 (23) *
Origin	n.s.	n.s.	n.s.
Sex	n.s.	-2.975 (31) **	-5.0077 (39) ***
INTERACTIONS			
Group (GD): phoneme	n.s.	**-2.558 (38) ***	n.s.
Group (SD): phoneme	n.s.	**n.s.**	n.s.
Group (GD): style	n.s.	n.s.	n.s.
Group (SD): style	1.758 (55) .	n.s.	n.s.
Group (GD): origin	-3.666 (46) ***	n.s.	n.s.
Group (SD): origin	-2.126 (46) *	n.s.	n.s.
Style: origin	n.s.	n.s.	n.s.
Style: phoneme	2.298 (100) *	n.s.	n.s.
Group (GD): sex	n.s.	n.s.	n.s.
Group (SD): sex	n.s.	n.s.	n.s.
Style: sex	n.s.	n.s.	n.s.
Origin: phoneme	n.s.	n.s.	n.s.
Origin: sex	n.s.	n.s.	n.s.
Phoneme: sex	n.s.	1.739 (28) .	2.109 (16) .
Group (GD): style: origin	2.694 (59) **	n.s.	n.s.
Group (SD): style: origin	n.s.	n.s.	n.s.
Group (GD): style: phoneme	n.s.	n.s.	1.885 (73) .
Group (SD): style: phoneme	n.s.	n.s.	n.s.
Group (GD): style: sex	n.s.	n.s.	n.s.
Group (SD): style: sex	n.s.	n.s.	n.s.
Style: origin: phoneme	n.s.	n.s.	n.s.
Style: origin: sex	n.s.	n.s.	n.s.
Origin: phoneme: sex	n.s.	n.s.	n.s.

Table 1. *Summary of the results of the regression models for Experiment 1. GD= Galician-dominant; SD= Spanish-dominant. Baselines for predictor variables: neofalante (group), text (style), rural (origin), front vowel or alveolar fricative /s/ (phoneme), female (sex). Numbers represent t-statistic and degrees of freedom (in brackets). *** p < 0.001 ** p < 0.01 * p < 0.05 . p < 0.10. Group effects in bold.*

The regression model (Table 1) demonstrated that there was a significant contrast between the *neofalantes* (M(raw)=0.414) and Galician-dominant (M(raw)=0.866) groups, but no significant contrast between *neofalantes* and Spanish-dominants (M(raw)=0.377). Figure 2 shows the Euclidean distance by group. Additionally, there was a main effect of speech style; overall, the Euclidean distance was significantly higher in the text (M(raw)=0.611) than in the wordlist (M(raw)=0.546).

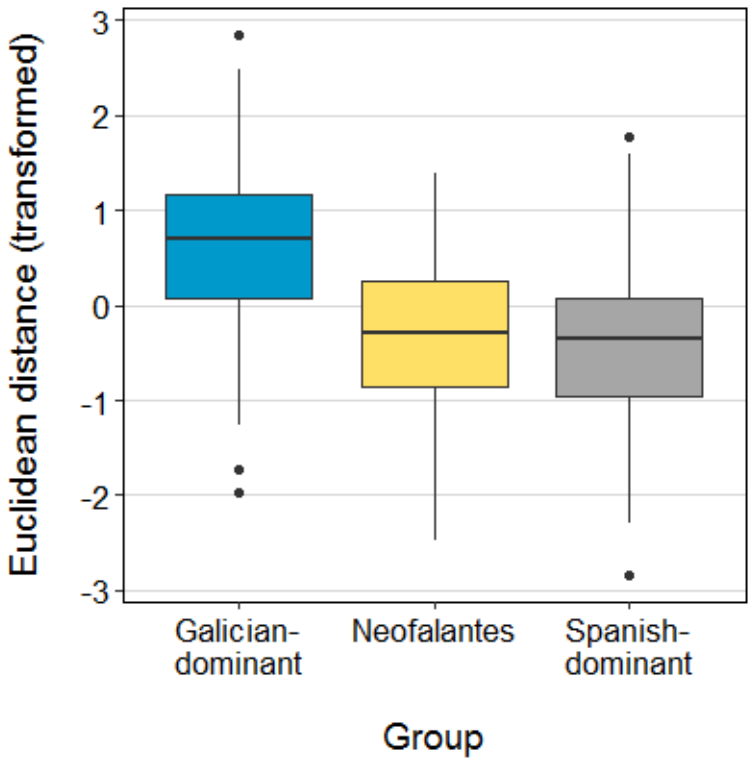

Figure 2: Boxplot showing the average Euclidean distance for the front and back mid-vowel contrasts by group (Galician-dominant, neofalantes, Spanish-dominant), averaged over vowel and speech style. Measurements were transformed using a rank-transformation to normality to correct for a skewed distribution.

The contrast between *neofalantes* and Galician-dominants was modulated by a significant interaction with origin, and so was the contrast between *neofalantes* and Spanish-dominants. The difference between Galician-dominants and *neofalantes* was bigger for those participants from rural settings, and much smaller for those from urban settings, such that urban Galician-dominant and *neofalantes* were very similar. For the

comparison between *neofalantes* and Spanish-dominants, the pattern was reversed: urban *neofalantes* had a higher Euclidean distance than the Spanish-dominant counterparts, but rural *neofalantes* had a lower Euclidean distance. Urban Galician-dominant speakers in this sample do not appear to produce a robust contrast between mid-vowels. However, given the limited sample size for the urban Galician-dominant group (N=5), this effect is difficult to interpret and needs replication with a larger sample.

The vowel-style interaction was significant, suggesting that the difference was larger for back vowels in the text (M(raw)=0.646) compared to the wordlist (M(raw)=0.472). This could be related to the fact that the vowels in the wordlist and text were not embedded in the same phonetic context. Finally, there was a three-way interaction, between group, style and origin for Galician-dominant vs *neofalantes*.

Voiceless sibilant fricatives

To investigate whether the three groups could produce the /s/-/ʃ/ contrast, CoG was used as the dependent variable in the model. Given that this measurement yields a skewed distribution, a rank-transformation to normality was applied to fit a linear mixed-effects regression model. The variables group, phoneme (alveolar /s/, post-alveolar /ʃ/), style, origin and sex were included as fixed factors in the model up to three-level interactions. All possible by-participant and by- word random slopes were included in the model (Barr et al., 2013).

As Table 1 shows, the regression model revealed a significant main effect of sex, as expected, suggesting that female speakers had a higher CoG overall when compared to male speakers. Although the effect of phoneme on its own was not significant, it was modulated by a significant interaction with the contrast between Galician-dominant

speakers and *neofalantes*, indicating that *neofalantes* behaved significantly differently from Galician-dominant but not Spanish-dominant speakers when producing /s/ and /ʃ/. Figure 3 shows CoG by phoneme by group. Given that /ʃ/ does not exist in Spanish and is not mentioned as present in impressionistic descriptions of Spanish-dominant accents of Galician (e.g., Regueira, 1999; González González, 2008), we had hypothesised that Spanish-dominants and potentially *neofalantes* might be unable to produce /ʃ/. However, there seem to be no clear differences in production between the three groups (Table 2). That said, there is more variation in /s/; Galician-dominant speakers seem to have a higher CoG, when compared to *neofalantes* and Spanish-dominants who do not differ from each other.

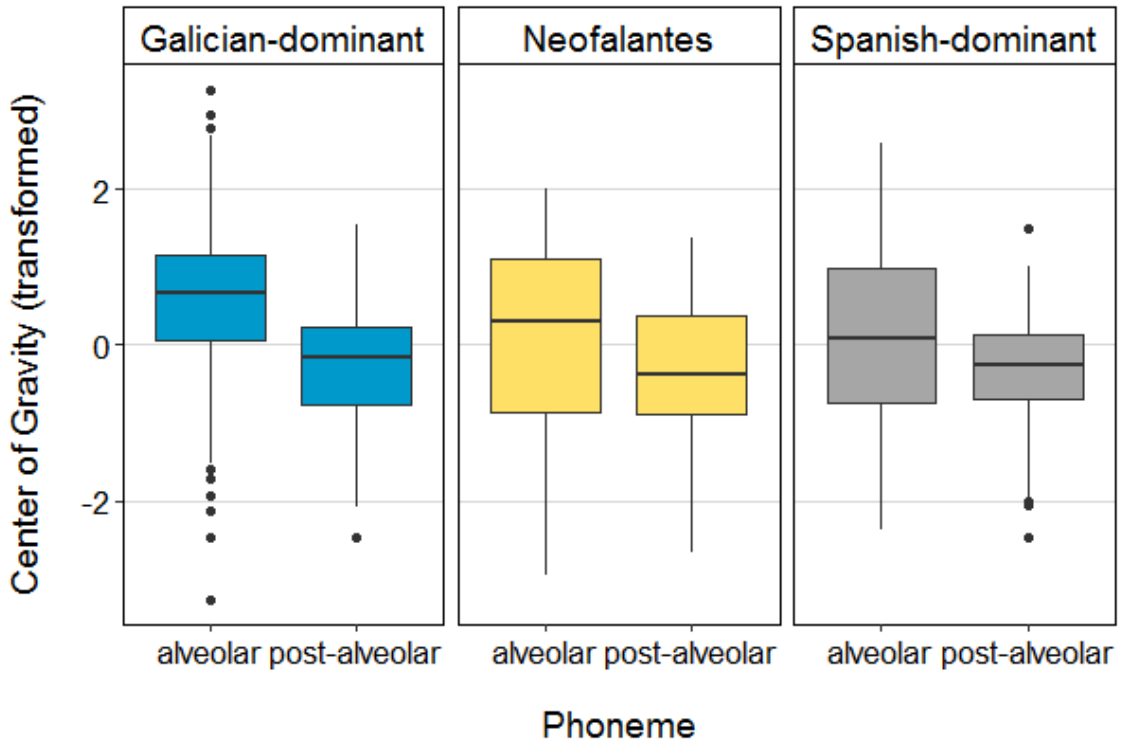

Figure 3: Boxplots showing transformed centre of gravity values for each phoneme (alveolar, post-alveolar) for the three groups (Galician-dominant, neofalantes, Spanish-dominant) averaged over speech style. Measurements were transformed using a rank-transformation to normality to correct for a skewed distribution.

Phoneme	Speaker group					
	GD female	GD male	*Neofalante* female	*Neofalante* female	SD female	SD male
/s/	5853	4656	5283	4084	5055	4125
/ʃ/	4346	3944	4247	3802	4064	3976

Table 2. *Mean centre of gravity (CoG) values in Hertz for each phoneme (alveolar* /s/*, postalveolar* /ʃ/*) for the three speaker groups, Galician-dominant (GD), Neofalantes and Spanish-dominant (SD), split by sex (female, male).*

To quantify the overlap between the two categories (i.e., /s/-/ʃ/) for each group, a logistic regression model was used as a classifier to predict the phoneme label from the transformed CoG. CoG was the only predictor included in the model. Higher prediction accuracy corresponds to less overlap in the distributions of CoG for the two fricatives. The model was scored as accurate if the probability of the true label for a given observation was above 50%. For Galician-dominants, the prediction accuracy was 71.3%, while for Spanish-dominants the accuracy was 56.2% and for *neofalantes* 58.4%. This indicates that the fricative categories were much more distinct for Galician-dominant speakers than for *neofalantes* and Spanish-dominants.

Unstressed word-final vowels

Only the mid unstressed vowels /e o/ were included in the analysis. Visual inspection of the data showed that for these vowels, reduction took place mainly along the F1 dimension, which is inversely correlated with vowel height (the higher F1, the lower the vowel) and so this measurement was used as the dependent variable. To compensate for the non-normality of the distribution a rank inverse normal transform was applied.

Preliminary observation of the data suggested that all groups had vowel reduction, but that although the differences were small, some had more reduction than others. A Welch two sample t-test showed that Galician-dominants' vowels were different from

those of Spanish-dominants (t=-2.4049(1263.8), p=.0163), confirming that the latter show less vowel reduction. To investigate whether *neofalantes* behaved like Galician-dominant or Spanish-dominant speakers, a linear mixed-effects regression model was fitted on the transformed F1 values. The variables of group, style, origin, vowel (front, back) and sex were included as fixed factors in the model up to three-level interactions. All possible by-participant and by-word random slopes were included in the model (Barr et al., 2013).

Table 1 summarises the output of the regression model. There was a significant contrast between the *neofalantes* and Spanish-dominant groups, but no significant contrast between the *neofalantes* and Galician-dominant groups. As displayed in Figure 4, Galician-dominants and *neofalantes* showed lower F1 values than Spanish-dominants, with *neofalantes* patterning more closely with Galician-dominants. The effect of style was significant, suggesting that vowels in the text had a lower F1 overall, that is, speakers tended to use more raised vowels overall. There was a highly significant effect of sex, due to male speakers having a lower F1, and therefore higher vowels. There was also a significant interaction between vowel and sex, indicating that the difference between male and female speakers was more pronounced for front vowels.

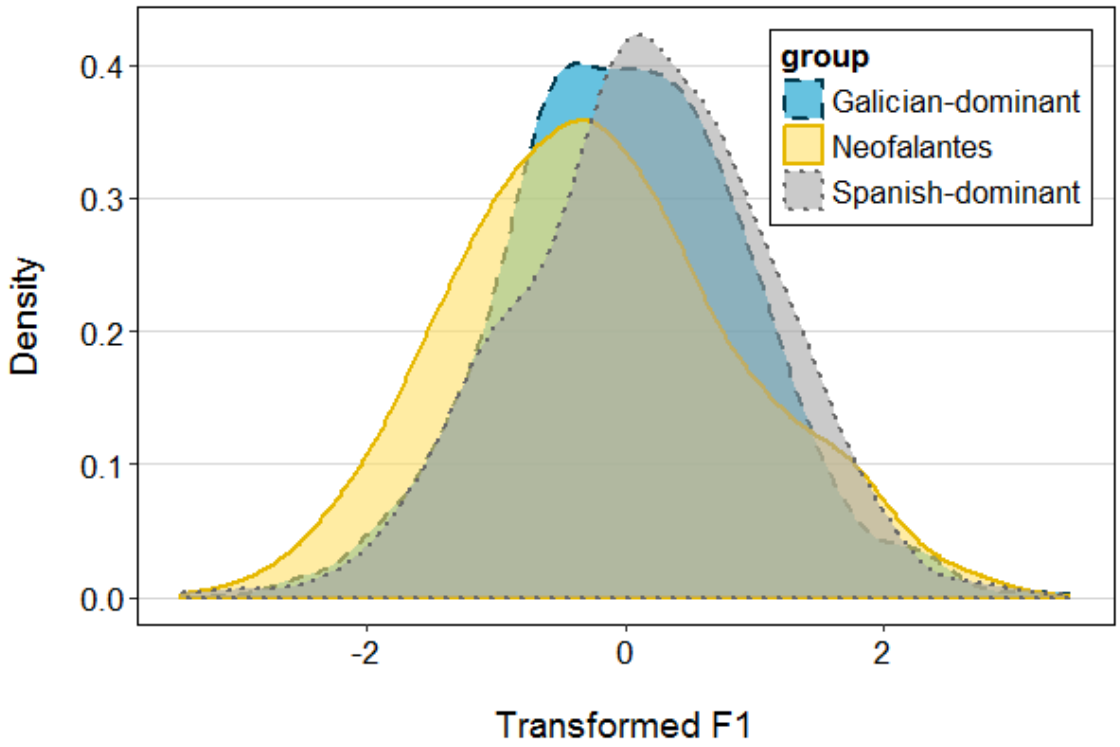

Figure 4: Density plot showing the transformed F1 distribution for the front and back unstressed word-final vowels, split by group (Galician-dominant, dashed line; neofalantes, solid line; Spanish-dominant, dotted line), averaged over vowel and speech style. Normalised F1 measurements were transformed using a rank-transformation to normality to correct for a skewed distribution.

2.3. ***Discussion***

There was little evidence to suggest that *neofalantes* had acquired the mid-vowel contrasts, patterning with Spanish-dominant speakers. However, *neofalantes* were able to acquire the fricative contrast, but so were Spanish-dominants, such that all speakers produced a /s/-/ʃ/ contrast. Surprisingly, all three groups were able to produce /ʃ/, which only exists in Galician, and there were no differences between groups for this phoneme. Nevertheless, there were differences in the magnitude of the contrast. Galician-dominants had a more distinct /s/-/ʃ/ contrast, while *neofalantes* and Spanish-dominants produced these phonemes with greater overlap. This difference was driven by differences in the production of /s/. Although there are no direct comparisons in the

literature, standard Castilian Spanish /s/ is often described as having an apical realisation and relatively low frequency values (cf. Martínez Celdrán & Fernández Planas, 2007). It is likely that the Galician realisation has a higher CoG, although different realisations have been found within Galician (cf. Labraña Barrero 2009; 2014; Regueira & Ginzo in press). It is therefore not surprising that in the current study Galician-dominants produce this phoneme differently from Spanish-dominant speakers. Overall then, the results indicate that although *neofalantes* are able to produce the fricative contrast, they do not change the way in which they do this after a switch in language dominance. Finally, although all groups used reduced vowels, the word-final vowel analysis suggested that *neofalantes* produced these more like Galician-dominant speakers.

In sum, there seem to be limits to what *neofalantes* can learn in terms of production. They are unable to acquire the mid-vowel contrasts, and do not change production of /s/ to match Galician-dominants. However, their accent is not exactly like that of Spanish-dominant bilinguals either; they produce unstressed word-final vowels more like Galician-dominant than Spanish-dominant speakers.

3. Experiment 2: Measurement of perception

Participants completed a vowel and a fricative identification task. Of interest was whether a change in language dominance would affect perception of the mid-vowel and sibilant fricative contrasts.

3.1. Method

Participants

Same as Experiment 1.

Materials

Participants completed two identification tasks. They identified naturally-produced words containing mid-vowels in stressed position (vowel identification task), and fricatives embedded in non-words on a synthetic continuum that ranged from /s/ to /ʃ/ (fricative identification task).

Vowel identification task. The stimuli consisted of the Galician minimal pairs *óso* [ˈɔsǫ] 'bone', *oso* [ˈosǫ] 'bear', *pé* [ˈpɛ] 'foot', *pe* [ˈpe] 'p', *só* [ˈsɔ] 'alone', *so* [ˈso] 'under', *té* [ˈtɛ] 'tea', *te* [ˈte] 't', embedded in the carrier sentence *Digo a palabra* ____ (I say the word ____). The carrier sentences were produced in two accents; (1) standard-accented Galician and (2) regionally-accented Galician. The latter included *gheada*, a very salient regional variant, which occurs in the Western half of Galicia, but is well-known in the whole territory and is often associated with 'traditional' Galician spoken in rural areas. By this process, [g] and [ɣ] are produced as [ɦ], [h], [x], [ɦ] or [ʕ], here giving [ˈdihǫapaˈlaβɾą] instead of [ˈdiɣǫapaˈlaβɾą]. This feature was included because it was hypothesised that it could act as a cue for the Galician-dominant accent. This manipulation only affected the carrier phrase. The same token of each target word was then spliced into the two carrier sentences, and thus the pronunciation of the target word did not vary between conditions. All stimuli were produced by the same male Galician-dominant speaker who was selected because he was able to produce both accents. Recordings were made in a sound attenuated room using a RODE NT1-A microphone directly connected to a PC via an Edirol processor with a sampling rate of 44.1kHz, 16-bit resolution. The speaker recorded two repetitions and then the best was selected for use in the experiment. Stimuli were band-pass filtered at 60-20,000Hz with a smoothing factor of 10. Finally, intensity was scaled to 70dB SPL and the files downsampled to 22,050Hz. All processing was carried out in Praat (Boersma & Weenink, 2014). Stimuli

were played over a laptop (ASUS A55V) via a Realtek HD Audio sound card, and were presented over headphones (Sennheiser HD 25-C II).

Fricative identification task. The stimuli were two-segment CV sequences that consisted of a fricative that varied in the place of articulation along a 22-step continuum from /s/ to /ʃ/ followed by the vowel /u/, giving the non-words "*su*" [ˈsu] or "*xu*" [ˈʃu] at the endpoints. These endpoints were based on natural tokens of /s/ and /ʃ/ recorded by the same Galician-dominant speaker as for the vowel identification task, with the recording procedure and processing also the same. The intermediate steps of the continuum were then created following the procedure described in Repp (1981) and McQueen (1991).

Briefly, the /s/ and /ʃ/ were excised from the natural recording from their onset to the zero crossing before the start of the vowel, and saved to individual wav files. The duration of each fricative was measured (/s/: 216.24ms; /ʃ/: 207.35ms) and the average duration calculated. The fricatives were then equalised for length in Praat (Boersma & Weenink, 2014) using PSOLA such that both were equal to the average duration, 211.91ms. These were used as the endpoints of the fricative continuum. The fricative portion of the intermediate stimuli was constructed by adding the amplitudes of the two waveforms in different proportions, giving 22 tokens each with a duration of 212ms. These were then spliced onto a natural token of /u/, excised from the recording of "*su*", creating 22 CV tokens where the fricative varied in equal steps from /s/ to /ʃ/. Lastly, intensity was scaled to 70dB and the files downsampled to 22,050Hz.

Four pilot participants, all Spanish-Galician bilinguals, completed the experiment to check the validity of the continuum.

Procedure

Participants completed the tasks in the same session as the production tasks (Experiment 1). Participants always completed the vowel identification task first. The vowel identification task consisted of 2 blocks (standard Galician, regional Galician) with the order of presentation counterbalanced across participants. In each block, participants identified the word they heard by clicking on the corresponding picture. In written Galician, open vowels can be signalled by an accent i.e., *óso* [ˈɔsǫ] (bone), *oso* [ˈosǫ] (bear), and so pictures were used to prevent orthographic cues influencing the results. Participants identified 4 repetitions of the 8 stimuli, giving a total of 32 trials per block. They heard each trial only once, with the order of presentation randomised across participants and the same stimulus never played twice in succession.

In the fricative identification task, participants identified whether they heard the non-word "*su*" [su] or "*xu*" [ʃu]. Before completing the task, participants completed a short practice session to familiarise them with the task. The practice included 10 different stimuli from the 22-step continuum presented in a randomised order. In the test block, participants identified 4 repetitions of the 22 stimuli, giving a total of 88 trials. They heard each trial only once, with the order of presentation randomised across participants and the same stimulus never played twice in a row.

3.2. Results

Mid vowels in stressed position

To investigate the effect of group on vowel identification, a mixed-effects logistic regression model was built with the binomial response (correct/incorrect) as the dependent variable, group (*neofalantes*, Galician-dominant, Spanish-dominant listeners), accent (standard, regional) and origin (urban, rural) as fixed factors and

participant and word as crossed random effects. Since word was included in the model as a random factor to account for the variance introduced by the different stimuli and there were two words per vowel, vowel was not included in the model as a fixed factor. Table 3 summarises the results of the model.

	Mid vowels	Sibilant fricatives
Intercept	7.481***	-8.500***
MAIN EFFECTS		
Group (GD)	**2.395 ***	-2.809 **
Group (SD)	**n.s.**	n.s.
Stimulus (c)	N/A	-30.652 ***
Origin	-1.746 .	n.s.
Accent	n.s.	N/A
INTERACTIONS		
Group (GD): stimulus (c)	N/A	**-2.835 ****
Group (SD): stimulus (c)	N/A	**n.s.**
Group (GD): origin	-2.771 **	n.s.
Group (SD): origin	n.s	2.056 *
Group (GD): accent	2.864 **	N/A
Group (SD): accent	2.087 *	N/A
Group (GD): stimulus (c): origin	N/A	n.s.
Group (SD): stimulus (c): origin	N/A	3.637 ***
Group (GD): accent: origin	n.s.	N/A
Group (SD): accent: origin	n.s.	N/A

Table 3. *Summary of the results of the regression models for Experiment 2. GD= Galician-dominant; SD= Spanish-dominant; c= centred; N/A: condition not included in the task. Baselines for categorical predictor variables: neofalante (group), rural (origin), standard (accent). Numbers represent Wald statistics (z-values). *** p < 0.001 ** p < 0.01 * p < 0.05 . p < 0.10. Group effects in bold.*

The model revealed a significant contrast between *neofalantes* (M=0.79) and Galician-dominant listeners, who performed at ceiling (M=0.92), but no significant contrast between *neofalantes* and Spanish-dominants (M=0.76). To test whether *neofalantes* and Spanish-dominants could identify the vowels above chance level, the dataset was compared to a random baseline. Two separate logistic regression analyses

for *neofalantes* and Spanish-dominants indicated that both groups performed significantly above chance.

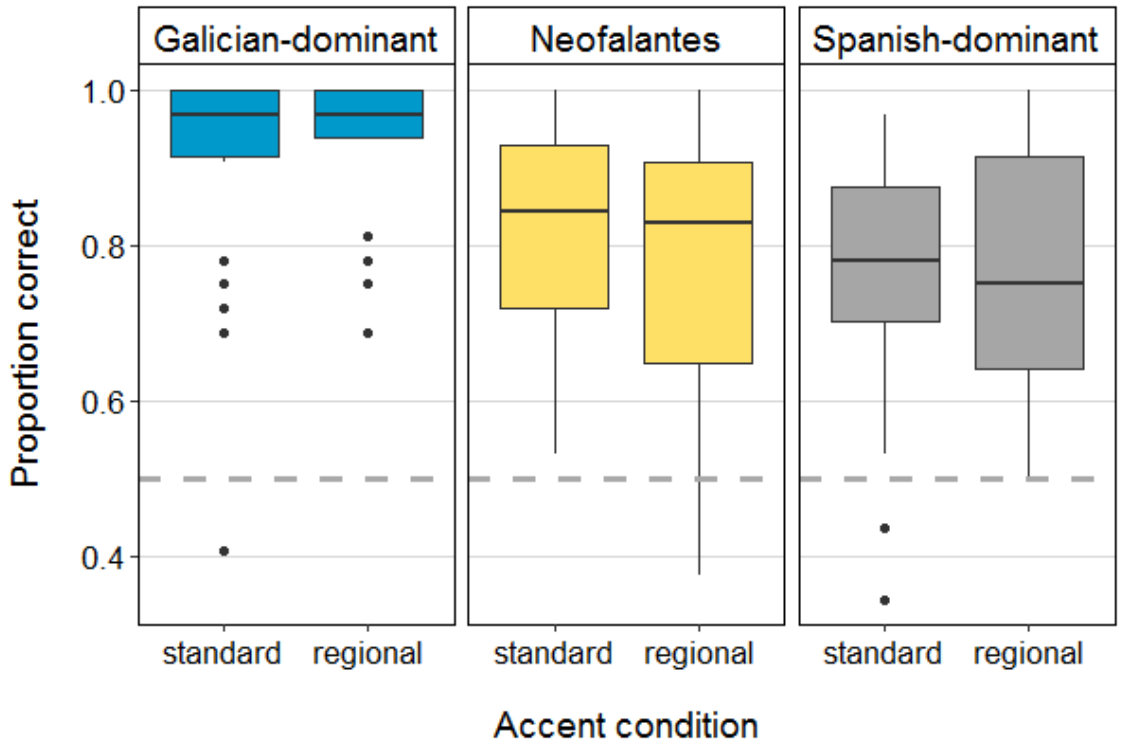

Figure 5: Boxplot showing vowel identification scores (proportion correct) by listener group (Galician-dominant, neofalantes, Spanish-dominant) and accent (standard Galician, regional Galician). The dashed line represents chance level performance.

In the main regression model, the contrast between *neofalantes* and Galician-dominants was modulated by a significant interaction with accent. This indicates that the difference between these two groups was smaller in the standard accent (M(GD)=0.90, M(NE)=0.82) than the regional accent condition (M(GD)=0.93, M(NE)=0.76). As displayed in Figure 5, *neofalantes* performed more poorly than the Galician-dominants overall, but did slightly better in the standard accent condition. The contrast between *neofalantes* and Spanish-dominants was also modulated by a significant interaction with accent. Although there was no overall difference in performance between these two groups, the difference in performance between

neofalantes and Spanish-dominants was greater in the standard condition (M(SD)=0.75, M(NE)=0.82) than the regional condition (M(SD)=0.76, M(NE)=0.76).

The main effect of origin approached significance and was modulated by a significant interaction with the contrast between Galician-dominants and *neofalantes*. This indicated that the difference between these groups was smaller for urban (M(GD)=0.79, M(NE)=0.81) than rural listeners (M(GD)=0.95, M(NE)=0.77), with urban Galician-dominant listeners performing more poorly than their rural counterparts, mirroring the production results.

Voiceless sibilant fricatives

To investigate the effect of group on the sibilant fricative continuum categorisation, a mixed-effects logistic regression model was fitted with the binomial response /s/-/ʃ/ as the dependent variable. The fixed factors included in the model were group (*neofalantes*, Galician-dominants, Spanish-dominants), stimulus (/s/-/ʃ/ continuum; centred) and origin (urban, rural), with participant as a crossed random effect.

Table 3 summarises the results of the model, which revealed a significant main effect of stimulus and a significant contrast between *neofalantes* and Galician-dominant listeners, but no significant contrast between *neofalantes* and Spanish-dominants. Moreover, the effect of origin was not significant. As expected, as the stimulus continuum increased from /s/ to /ʃ/, the less likely it was for listeners to choose /s/. In terms of group, Galician-dominants chose /s/ less than *neofalantes* overall. The contrast between the *neofalantes* and Galician-dominant groups was modulated by a significant interaction with phoneme, suggesting that *neofalantes*' categorisation of the phonemes was different from that of Galician-dominant listeners, but not from that of Spanish-dominants. Figure 6 shows the identification scores and the model predictions for the

three groups and suggests that, although all three groups have categorical perception of this contrast, Galician-dominant listeners start categorising the stimuli as /ʃ/ earlier than the two other groups.

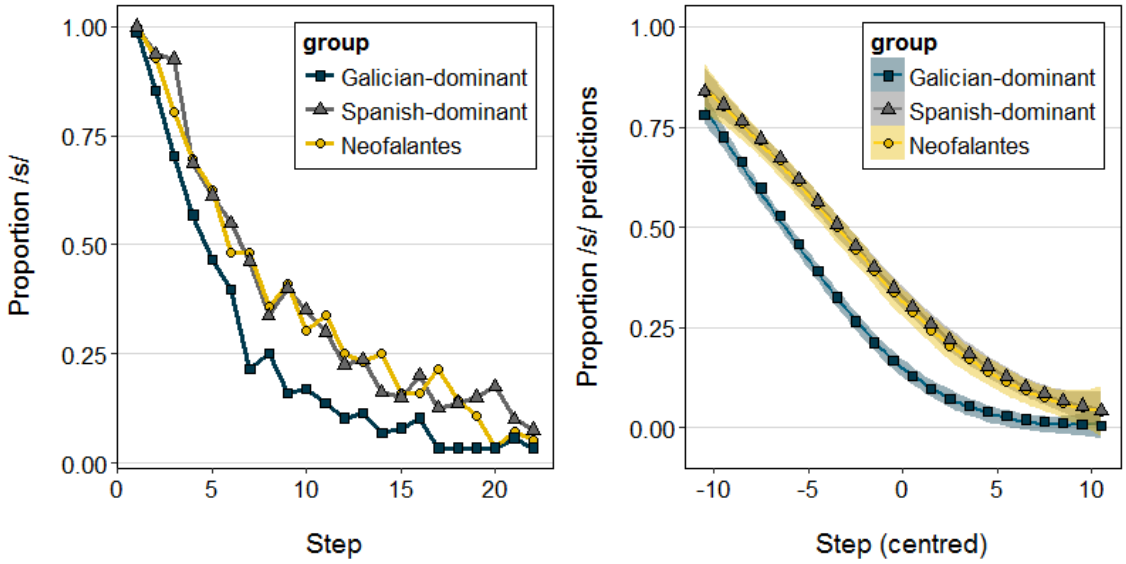

Figure 6: Raw proportion of /s/ response according to stimulus step (1-22) by group (Galician-dominant, squares; Spanish-dominant, triangles; neofalantes, circles; left panel) and model predictions according to centred stimulus step by group (right panel).

Additionally, the contrast between *neofalantes* and Spanish-dominant listeners was modulated by a significant interaction with origin; the frequency of choosing /s/ was different for urban and rural participants in these two groups. This effect was further modulated by a significant three-way interaction with stimulus, which also indicated that there were differences between these two groups when taking stimulus into account; urban *neofalantes* had an earlier categorisation boundary than rural *neofalantes*, whereas Spanish-dominants showed the opposite pattern.

3.3. Discussion

Overall, *neofalantes* did not differ in their perception from Spanish-dominants, indicating that they had not changed to behave more like Galician-dominants. Although

both *neofalantes* and Spanish-dominants performed relatively well with the mid-vowel contrasts, they performed more poorly than Galician-dominants. In addition, *neofalantes* behaved differently from both groups in the different accent conditions; their perception of the mid-vowels when listening to the standard accented stimuli was slightly closer to that of Galician-dominants and better than that of Spanish-dominants. For fricatives, *neofalantes* likewise patterned with Spanish-dominant listeners. Both groups could perceive the /s/-/ʃ/ contrast, but they had a later phoneme categorisation boundary when compared to Galician-dominants, i.e., they still heard /s/ for tokens where Galician-dominants already heard /ʃ/. This mirrors the production results that showed that *neofalantes* and Spanish-dominants have a lower CoG for /s/ than Galician-dominants.

4. General discussion

This study investigated whether *neofalantes* changed aspects of their speech production and perception after switching language dominance to Galician. Change was inferred by comparing this group to two control groups, Galician-dominant and Spanish-dominant bilinguals. If they patterned with Spanish-dominants when producing and perceiving variables that are specific to Galician, it was assumed that there was no change, while if they patterned with Galician-dominants it was assumed that there was a change. In production, there was little evidence to suggest that *neofalantes* were able to produce the two mid-vowel contrasts that do not exist in Spanish, behaving like Spanish-dominant speakers. They were able to produce a sibilant consonant contrast that does not exist in Spanish, but so were Spanish-dominants. However, both *neofalantes* and Spanish-dominants produced this contrast differently from Galician-dominants, who had a more distinct contrast. Word-final vowels, on the other hand, which are a highly salient feature of Galician, were produced by *neofalantes* with a more Galician-like realisation, different from Spanish-dominants. In perception,

neofalantes behaved like Spanish-dominants for both mid-vowel and the fricative contrasts. Both groups were able to identify the mid-vowel contrast, but their performance was poorer than that of Galician-dominants. Likewise, both groups were able to identify the fricative contrast, but had a different category boundary from Galician-dominants.

Previous research with Spanish-Catalan bilinguals has shown that early exposure is not enough for dominant bilinguals to acquire native-like categories in their non-dominant language, and this has been attributed to a lack of plasticity (Pallier et al., 1997; Sebastián-Gallés & Soto-Faraco, 1999; Pallier et al., 2001). Likewise, *neofalantes* had limited success in acquiring the front and back mid-vowel contrast in production and did not perform like Galician-dominants in perception. However, it has also been argued that difficulties in L2 perception are due to continued use of the L1 (Flege & Mackay, 2004; Mora et al., 2011, 2015). Given that the bilinguals in the Catalan studies continued to use their dominant language (in this case, Spanish), one could hypothesise instead that the failure to establish native-like phonetic categories was because the L1 continued to exert a strong influence on perception and therefore, the L2. Nevertheless, our results show that even with extensive use of the L2 and a high motivation to learn, dominant bilinguals are unable to form new, native-like phonetic categories in production or perception when they switch late in life. It seems more likely then that *neofalantes* process their new, dominant language through their former dominant language categories.

Theories of cross language speech perception such as PAM/PAM-L2 (Best, 1994, 1995; Best & Tyler, 2007) and the SLM (Flege, 1992, 1995), have proposed that certain phonetic contrasts are more difficult to perceive than others and that this leads to difficulties in production. According to these models, the difficulty can be predicted by

the phonetic similarities of the first and second languages. The contrast between open and close mid-vowels is a difficult one for *neofalantes* (and Spanish-dominants), because the Galician contrasts are both a good match to the single Spanish categories. However, although their category boundary was different from Galician-dominant listeners, both *neofalantes* and Spanish-dominants were able to perceive and produce the fricative contrast that does not exist in Spanish. Flege (1995) postulates that bilinguals are able to establish a new phonetic category for an L2 sound that differs phonetically from the closest L1 sound if they are able to discern at least some of the phonetic differences between the L1 and L2 sounds. One possibility is that this contrast is more acoustically distinct than the mid-vowel contrasts, such that both Spanish-dominants and *neofalantes* are able to establish a category even though this does not match that of native speakers.

In contrast, word-final vowels seem to be more mutable. Although *neofalantes* behaved like Spanish-dominants in their production of mid-vowels and fricatives, they produced word-final vowels like Galician-dominants; all speakers used reduced vowels, but *neofalantes* patterned with Galician-dominants in having a greater amount of reduction than Spanish-dominants. Although it is possible that word-final reduction may be a more predictable pattern and that this facilitates acquisition, word-final vowels are a highly salient characteristic of the Galician accent. It is thus possible that social factors also played a role in production of this variable. *Neofalantes* switch language dominance for ideological reasons, and when they do, they are often very aware that they do not speak like Galician-dominants. They are very motivated to "learn" the language and most of our participants reported having made a conscious effort to improve their pronunciation, i.e., to speak with a more native-like accent. One interpretation is that *neofalantes* use this feature, whether consciously or

subconsciously, to fit in with their new group of Galician-dominant speakers. This is similar to findings from studies of accent change within the same language; Evans and Iverson (2007) showed that speakers who changed their accent late in life (young adulthood) acquire some, but not all the phonetic features that characterise their new accent. However, their realisation was not like that of native speakers, and not all the speakers showed the same changes in production. These individual differences were interpreted as reflecting the way speakers chose to present themselves to the world. In a bilingual context, Amengual (2015) found that Spanish-dominant bilinguals did not differ from Catalan-dominant bilinguals in their production of some reduced vowels in Majorcan Catalan; both groups produced /a/ as a reduced centralised [ə] in unstressed position. This was interpreted as being a result of the "construction of socio-indexical phonological categories based on a stronger identification with the prestigious Standard Catalan variety" (2015, p. 4). In the Galician community, although the reduced vowels might not be associated with the prestigious variety, they are indeed associated with a Galician-like accent, thus meaning that they could be used to signal Galician identity.

Studies in the lab have also shown that there is flexibility in production and perception in adulthood. For example, high variability phonetic laboratory training studies have shown that L2 listeners can improve in their identification of phonetic contrasts that do not exist in their L1 (Logan, Lively & Pisoni, 1991; Lively, Logan & Pisoni, 1993; Iverson & Evans, 2009), that this knowledge can be transferred to production (Bradlow, Pisoni, Akahane-Yamada & Tohkura, 1997) and that it is retained after a few months (Bradlow, Akahane-Yamada, Pisoni & Tohkura, 1999). However, there appear to be limits to this such that even early exposure to an L2 in a bilingual environment is not enough to acquire native-like categories in the non-dominant language (Pallier et al., 1997). The current study is in line with these findings and

provides further evidence that "real life training" or in this case, extensive naturalistic exposure to and use of the L2, is not enough for dominant bilinguals to acquire native-like categories in their non-dominant language. Even with what could be seen as ideal circumstances for learning - early and extensive exposure, almost exclusive use of the L2 and very high motivation - L2 production and perception still seem to be filtered by L1 categories. One possibility is that underlying categories are very difficult to change, and that although, with experience, individuals can improve at mapping new categories onto native ones, they do not create new categories (Iverson & Evans, 2009). The focus of this study is group differences, and it is relevant to highlight that individual differences (e.g., learning ability or other cognitive skills) might play a role in the acquisition of such phonetic contrasts. That is, it is perhaps not the case that no *neofalante* can ever learn Galician-like contrasts, but this at least seems very difficult.

These results thus argue for a central role of early exposure in phonetic processing. Although studies of Korean adoptees adopted by French families and exposed exclusively to French from between the ages of 2 and 9 years old (Pallier, Dehaene, Poline, LeBihan, Argenti, Dupoux & Melher, 2003; Ventureyra, Pallier, Yoo, 2004), have indicated that all traces of attunement to the L1 sound system are lost by adulthood, new research with Chinese adoptees in Canada, also exposed exclusively to French since adoption, has shown that early experience can have lasting effects (Pierce, Chen, Delcenserie, Genesee & Klein, 2015). Although their performance on behavioural tasks did not differ from that of French monolinguals, Chinese adoptees' brain activation patterns were more similar to those of Chinese-French bilinguals. This suggests that early exposure to a language continues to influence the neural processing of a subsequently learned language sounds years later, even in highly proficient, early-exposed users.

In conclusion, these findings suggest that native-like production and perception of new phonetic contrasts is difficult to attain. Despite early exposure, extensive use and high motivation, there was little evidence to indicate that *neofalantes* acquired the Galician mid-vowel contrasts in production and perception, and they behaved more similarly to Spanish-dominants in their production and perception of the fricative contrast. However, they produced unstressed word-final vowels in the same way as Galician-dominants. Together, this results in a hybrid variety different from that used by Galician- and Spanish-dominants, and characterised by the effects of a long-term switch in language dominance. Although underlying category representations thus appear hard to change, with modifications to production and perception constrained by early experience with a particular language, the resulting hybrid categories may function as opportunities to mark identity within a particular community.

5. References

Adank, P., Smits, R., & van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustical Society of America*, *116*, 3099–3107.

Amengual, M. (2014). The perception and production of language-specific mid-vowel contrasts: Shifting the focus to the bilingual individual in early language input conditions. *International Journal of Bilingualism*, 1–20.

Amengual, M. (2015). The acoustic realization of the /a/-/ə/ alternation in Majorcan Catalan. In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*. Glasgow, Scotland: University of Glasgow.

Amengual, M. (2016). Cross-Linguistic Influence in the Bilingual Mental Lexicon: Evidence of Cognate Effects in the Phonetic Production and Processing of a Vowel Contrast. *Frontiers in Psychology, 7,* 1–17.

Amengual, M., & Chamorro, P. (2015). The Effects of Language Dominance in the Perception and Production of the Galician Mid Vowel Contrasts. *Phonetica, 72,* 207–236.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: a perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (eds.), *The development of speech perception: the transition from speech sounds to spoken words*, pp. 167–224. Cambridge, MA: MIT Press.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, pp. 171–206. Timonium, MD: York Press.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S. Bohn (eds.), *Second language speech learning: The role of language experience in speech perception and production*, pp. 13–34. Amsterdam: John Benjamins.

Boersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer [Computer program]. Version 5.3.59, retrieved 20 November 2013 from http://www.praat.org/

Bosch, L., Costa, A., & Sebastián-Gallés, N. (2000). First and second language vowel perception in early bilinguals. *European Journal of Cognitive Psychology, 12,* 189–221.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English/r/and/l: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America, 101,* 2299–2310.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: long-term retention of learning in perception and production. *Perception & Psychophysics, 61,* 977–985.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1989). Limits on bilingualism. *Nature, 340,* 229–230.

Drummond, R. (2012a). Aspects of identity in a second language: ING variation in the speech of Polish migrants living in Manchester, UK. *Language Variation and Change, 24,* 107–133.

Drummond, R. (2012b). The Manchester Polish STRUT: Dialect Acquisition in a Second Language. *Journal of English Linguistics, XX(X),* 1–29.

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2010). Limits on bilingualism revisited: Stress "deafness" in simultaneous French-Spanish bilinguals. *Cognition, 114,* 266–275.

Eckert, P. (2000). *Linguistic variation as social practice*. Oxford: Blackwell.

Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics, 12,* 453–476.

Evans, B. G., & Iverson, P. (2004). Vowel normalization for accent: An investigation of best exemplar locations in northern and southern British English sentences. *The Journal of the Acoustical Society of America, 115,* 352.

Evans, B. G., & Iverson, P. (2007). Plasticity in vowel perception and production: a study of accent change in young adults. *The Journal of the Acoustical Society of America, 121,* 3814–26.

Flege, J. E. (1995). Second Language Speech Learning: Theory, Findings, and Problems. In W. Strange (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, pp. 233–277. Timonium, MD: York Press.

Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America, 106,* 2973–2987.

Flege, J. E., & MacKay, I. R. A. (2004). Perceiving Vowels in a Second Language. *Studies in Second Language Acquisition, 26,* 1–34.

Foulkes, P., & Docherty, G. J. (2006). The social life of phonetics and phonology. *Journal of Phonetics, 34,* 409–438.

García-Mateo, C., Cardenal, A., Regueira Fernández, X. L., Fernández Rei, E., Martínez, M., Seara, R., Varela, R., & Basanta Llanes, N. (2014). CORILGA: a Galician Multilevel Annotated Speech Corpus for Linguistic Analysis. *9th Language Resources and Evaluation Conference Reykjavik.*

González González, M. (2008). O novo galego urbano. In M. Brea, F. Fernández Rei, & X. L. Regueira (eds.), *Cada palabra pesaba, cada palabra medía. Homenaxe a Antón Santamarina*, pp. 363–374. Santiago de Compostela: Universidade de Santiago de Compostela.

Hazan, V., & Boulakia, G. (1993). Perception and Production of a Voicing Contrast by French-English Bilinguals. *Language and Speech, 36,* 17–38.

IGE [Instituto Galego de Estatística] (2008): Enquisa estrutural a fogares. Coñecemento e uso do galego 2008. Santiago: Instituto Galego de Estatística.

IGE [Instituto Galego de Estatística] (2013): Enquisa estrutural a fogares. Coñecemento e uso do galego 2013. Santiago: Instituto Galego de Estatística.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., & Diesch, E. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87,* 47–57.

Iverson, P., & Evans, B. G. (2007). Learning English vowels with different first-language vowel systems: perception of formant targets, formant movement, and duration. *The Journal of the Acoustical Society of America, 122,* 2842–2854.

Iverson, P., & Evans, B. G. (2009). Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *The Journal of the Acoustical Society of America, 126,* 866–77.

Jaffe, A. (2015). Defining the new speaker: theoretical perspectives and learner trajectories. *International Journal of the Sociology of Language, 231,* 21–44.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America, 108,* 1252–1263.

Kuznetsova, A., Bruun Brockhoff, P., & Haubo Bojesen Christensen, R. (2016). lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-32. https://CRAN.R-project.org/package=lmerTest

Labraña Barrero, S. (2009). Las consonantes fricativas de la lengua gallega. *Estudios de Fonética Experimental, 18,* 193-213.

Labraña Barrero, S. (2014). Parámetros acústicos dos sons fricativos da lingua galega. *Estudios de Fonética Experimental, 23,* 203-243.

LimeSurvey Project Team / Carsten Schmitz. (2012). LimeSurvey: An Open Source survey tool. Hamburg, Germany. Retrieved from www.limesurvey.org

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America, 94,* 1242–1255.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America, 89,* 874–886.

McQueen, J. M. (1991). The influence of the lexicon on phonetic categorization: stimulus quality in word-final ambiguity. *Journal of Experimental Psychology. Human Perception and Performance, 17,* 433–443.

Martínez Celdrán, E., & Fernández Planas, A. M. (2007). *Manual de fonética española. Articulaciones y sonidos del español*. Barcelona: Ariel.

Molinos Castro, R. (2002). As vocais átonas finais en galego: estudio acústico. *Cadernos de Lingua, 24,* 55-91.

Mora, J. C., Keidel, J. L., & Flege, J. E. (2011). Why are the Catalan contrasts between /e/-/ɛ/ and /o/-/ɔ/ so difficult for even early Spanish-Catalan bilinguals to perceive? In M. Wrembel, M. Kul & K. Dziubalska-Kołaczyk (eds.), *Achievements and*

perspectives in the acquisition of second language speech: New Sounds 2010, pp. 183–193. Bern, Switzerland: Peter Lang.

Mora, J. C., Keidel, J. L., & Flege, J. E. (2015). Effects of Spanish use on the production of Catalan vowels by early Spanish-Catalan bilinguals. In J. Romero & M. Riero (eds.), *The Phonetics-Phonology Interface. Representations and methodologies*, pp. 33–53. Amsterdam: John Benjamins.

Nance, C. (2015). "New" Scottish Gaelic speakers in Glasgow: A phonetic study of language revitalisation. *Language in Society, 44,* 553–579.

Nance, C., McLeod, W., O'Rourke, B., & Dunmore, S. (2016). Identity, Accent Aim, and Motivation in Second Language Users: New Scottish Gaelic Speakers' Use of Phonetic Variation. *Journal of Sociolinguistics, 20,* 164–191.

O'Rourke, B., & Ramallo, F. (2011). The native-non-native dichotomy in minority language contexts: Comparisons between Irish and Galician. *Language Problems and Language Planning, 35,* 139–159.

O'Rourke, B., & Ramallo, F. (2013a). Competing ideologies of linguistic authority amongst new speakers in contemporary Galicia. *Language in Society, 42*(3), 287–305.

O'Rourke, B., & Ramallo, F. (2013b). "A miña variedade é defectuosa": a lexitimidade social das neofalas. *Estudios de Linguistica Galega, 5,* 89–103.

O'Rourke, B., & Ramallo, F. (2015). Neofalantes as an active minority: Understanding language practices and motivations for change amongst new speakers of Galician. *International Journal of the Sociology of Language, 231,* 147–165.

O'Rourke, B., Pujolar, J., & Ramallo, F. (2015). New speakers of minority languages: the challenging opportunity–Foreword. *International Journal of the Sociology of Language, 231*, 1–20.

Pallier, C., Bosch, L., & Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition, 64*, B9-17.

Pallier, C., Dehaene, S., Poline, J.-B., Argenti, A.-M., Dupoux, E., & Mehler, J. (2003). Brain imaging of language plasticity: can a second language replace the first? *Cerebral Cortex*, 155–161.

Pierce, L. J., Chen, J.-K., Delcenserie, A., Genesee, F., & Klein, D. (2015). Past experience shapes ongoing neural patterns for language. *Nature Communications, 6*, 1-11.

Pujolar, J., & Puigdevall, M. (2015). Linguistic mudes: how to become a new speaker in Catalonia. *International Journal of the Sociology of Language, 231*, 167–187.

R Core Team (2016). R: A language and environment for statistical computing. Version 3.2.4. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Ramallo, F. (2013). Neofalantismo. In E. Gugenberger, H. Monteagudo, & G. Rei-Doval (eds.), *Contacto de linguas, hibrididade, cambio: contextos, procesos e consecuencias*, pp. 245–258. Santiago de Compostela: Consello da Cultura Galega/Instituto da Lingua Galega.

Ramallo, F., & O'Rourke, B. (2014). Perfiles de neohablantes de gallego. *Digithum, 16*, 98–105.

Ramon-Casas, M., Swingley, D., Sebastián-Gallés, N., & Bosch, L. (2010). Vowel categorization during word recognition in bilingual toddlers. *Cognitive Psychology, 59,* 96–121.

Real Academia Galega (ed.) (2009). *Mapa sociolingüístico de Galicia 2004. Vol. II: Usos lingüísticos en Galicia*. A Coruña: RAG.

Regueira, X. L. (1999). Estándar oral e variación social da lingua galega. In R. Álvarez & D. Vilavedra (eds.), *Cinguidos por unha arela común: homenaxe ó profesor Xesús Alonso Montero*, pp. 855–875. Santiago de Compostela: Universidade.

Regueira, X. L. (1994). Modelos fonéticos e autenticidade lingüística. *Cadernos de Lingua, 10,* 37-60.

Regueira, X. L. (2007). Vocais finais en galego e en portugués: un estudio acústico. *Actas do VII Congreso Internacional de Estudos Galegos*, pp. 859–876.

Regueira, X. L. (2012). Autenticidade e calidade da lingua, purismo e planificación lingüística no galego actual. *Estudos de Lingüística Galega, 4*, 187–201.

Regueira, X. L., & Ginzo, M. J. (in press). A cross-linguistic study of voiceless fricative sibilants in Galician and European Portuguese. In M. Gibson & J. Gil (eds.). *Romance Phonetics and Phonology*. Oxford: Oxford University Press.

Repp, B. H. (1981). Two strategies in fricative discrimination. *Perception & Psychophysics, 30*, 217–227.

Rojo, G. (2004). El español de Galicia. In R. Cano (ed.), *Historia de la lengua española*, pp. 1087–1101. Barcelona: Ariel.

Sebastián-Gallés, N., & Soto-Faraco, S. (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition, 72,* 111–123.

Sebastián-Gallés, N., Echeverría, S., & Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language, 52,* 240–255.

Sebastián-Gallés, N., & Díaz, B. (2012). First and Second Language Speech Perception: Graded Learning. *Language Learning*, *62,* 131–147.

Simonet, M. (2010). Dark and clear laterals in Catalan and Spanish: Interaction of phonetic categories in early bilinguals. *Journal of Phonetics, 38,* 663–678.

Simonet, M. (2011). Production of a Catalan-specific vowel contrast by early Spanish-Catalan bilinguals. *Phonetica, 68,* 88–110.

Ventureyra, V. A. G., Pallier, C., & Yoo, H. Y. (2004). The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics, 17, 7*9–91.

Appendix 1: Language background questionnaire

This section presents the questions from the language background questionnaire. It should be noted that this representation corresponds to an English translation of the actual questionnaire, which was presented in Galician on LimeSurvey (2012).

General information

1. Name:
2. Sex: ☐ Female ☐ Male
3. Address
4. Email address
5. Mobile number
6. Date of birth
7. Place of birth
8. Place of birth mother
9. Place of birth father
10. Place of residence during the academic year
11. Place of residence during the rest of the year
12. Have you lived somewhere else?

    ☐ Yes

    ☐ No (go to question 14)

13. Write down the place and dates

Place	From	To

14. Use this space if you would like to make a comment

15. How old were you when you learnt Spanish? How did you learn it?

16. How old were you when you learnt Galician? How did you learn it?

17. Select the option that best describes your linguistic background

    a. The language I use the most is Spanish
    b. The language I use the most is Galician
    c. I use both languages equally

If participant clicked a. The language I use the most is Spanish

i. Have you always spoken Spanish?

☐ Yes, I have

☐ No, I used to speak Galician more and now I speak Spanish more

ii. How old were you when you started speaking Spanish?

iii. Why did you switch languages?

If participant clicked b. The language I use the most is Galician

i. Have you always spoken Galician?

☐ Yes, I have (go to question 18)

☐ No, I used to speak Galician more and now I speak Spanish more

ii. How old were you when you started speaking Galician?

iii. Why did you switch languages?

iv. Have you made any effort to improve the way you speak Galician?

☐ Yes

How?

☐ No

v. Do you think it is important to have a correct pronunciation in Galician?

☐ Yes

Why?

☐ No

Why not?

☐ I haven't thought about it

vi. Have you made any effort to improve your pronunciation in Galician?

☐ Yes

How?

☐ No

If participant clicked c. I use both languages equally

i. Have you always spoken both languages equally?

☐ Yes

☐ No, I used to speak more Galician

☐ No, I used to speak more Spanish

18. Use this space if you would like to make a comment

19. Language in education

	Galician	More Galician than Spanish	Both	More Spanish than Galician	Spanish
Primary Education	○	○	○	○	○
Secondary Education	○	◉	○	○	○
University	○	○	○	◉	○

20. Which do(es) ---------- speak the most?

	Galician	More Galician than Spanish	Both	More Spanish than Galician	Spanish
your mother	◉	○	○	○	○
your father	○	◉	○	○	○
your siblings	○	○	○	◉	○

21. Which language do(es) ---------- speak to you?

	Galician	More Galician than Spanish	Both	More Spanish than Galician	Spanish
your mother	○	○	○	○	○
your father	○	◉	○	○	○
your siblings	○	○	○	◉	○

22. Do you speak other languages?

☐ Yes

☐ No (go to question 24)

23. Indicate which languages you speak, the age you started learning them and your competence level

Language	Age of acquisition	Competence

24. What do you study?

Degree:

Year:

25. Do you have any hearing impairment?

☐ Yes
Which one?
☐ No

26. Do you have any speech or language disorder?

☐ Yes
Which one?
☐ No

27. Would you be willing to participate in a further study?

☐ Yes

☐ No

28. Use this space if you would like to make a comment

29. Do you belong to any of these organizations?

- Sports organizations ☐ Yes ☐ No (if yes, which ones?)
- ONGs ☐ Yes ☐ No (if yes, which ones?)
- Organizations in favour of Spanish ☐ Yes ☐ No (if yes, which ones?)
- Organizations in favour of Galician ☐ Yes ☐ No (if yes, which ones?)
- Student organizations ☐ Yes ☐ No (if yes, which ones?)
- Political organizations ☐ Yes ☐ No (if yes, which ones?)

30. Use this space if you would like to make a comment

Language use

31. Which language do you speak…

	Galician	More Galician than Spanish	Both	More Spanish than Galician	Spanish	Not applicable
to your mother?	○	○	○	○	○	○
to your father?	○	○	○	○	○	○
to your siblings?	○	○	○	○	○	○
to your partner?	○	○	○	○	○	○
to your maternal grandparents?	○	○	○	○	○	○
to your paternal grandparents?	○	○	○	○	○	○
to your closest friends?	○	○	○	○	○	○
to your classmates?	○	○	○	○	○	○
to your lecturers?	○	○	○	○	○	○
to the doctor?	○	○	○	○	○	○
to strangers?	○	○	○	○	○	○
at work?	○	○	○	○	○	○
when shopping?	○	○	○	○	○	○
when flirting?	○	○	○	○	○	○

32. In which language…

		More Galician than Spanish	Both	More Spanish than Galician	Spanish	Not applicable
do you dream?	○	○	○	○	○	○
do you think?	○	○	○	○	○	○
do you count?	○	○	○	○	○	○
do you swear?	○	○	○	○	○	○
do you tell jokes?	○	○	○	○	○	○

do you take notes?						
do you write forms/formal letters (bureocracy)?						
do you use in social networks (Facebook, Twitter…)?						

33. Use this space if you would like to make a comment

Appendix B: Materials

Experiment 1 wordlist

Table 4. *List of target words included in the wordlist and used for analysis in Experiment 1. Each of these words was recorded in phrase-final position in the carrier sentence* digo a palabra _____ (I say the word _____) *and in phrase-medial position in the carrier sentence* digo a palabra _____ con coidado (I say the word _____ carefully). *For the mid-vowel set, only the mid-vowels* /e ɛ o ɔ / *were included in the statistical analysis. The vowels* /a i u/ *were used in the normalisation procedure and were included in plots for reference. For the word-final vowel set, only the mid unstressed vowels were included in the statistical analysis;* [ạ] *was included for reference.* F *= feminine gender,* N *= noun,* V *= verb.*

Variable	Segment	Word	Transcription	English translation
Reference vowel	/a/	*pazo*	[ˈpaθo̰]	pazo
Reference vowel	/i/	*pita*	[ˈpitạ]	hen
Reference vowel	/u/	*pucho*	[ˈpu͡tʃo̰]	calf
Mid vowel	/ɛ/	*peza*	[ˈpɛθạ]	piece
Mid vowel	/e/	*peto*	[ˈpeto̰]	pocket
Mid vowel	/e/	*seca*	[ˈsekạ]	dry (F)
Mid vowel	/ɔ/	*pote*	[ˈpɔtẹ]	pot
Mid vowel	/ɔ/	*sota*	[ˈsɔtạ]	knave (cards)
Mid vowel	/o/	*pozo*	[ˈpoθo̰]	well (N)
Mid vowel	/o/	*sopa*	[ˈsopạ]	soup
Fricative	/s/	*pase*	[ˈpasẹ]	pass
Fricative	/ʃ/	*paxe*	[ˈpaʃẹ]	page
Reference vowel	[ạ]	*peza*	[ˈpɛθạ]	piece
Reference vowel	[ạ]	*sopa*	[ˈsopạ]	soup
Reference vowel	[ạ]	*pata*	[ˈpatạ]	paw
Reference vowel	[ạ]	*pita*	[ˈpitạ]	hen
Reference vowel	[ạ]	*seca*	[ˈsekạ]	dry (F)
Reference vowel	[ạ]	*sota*	[ˈsɔtạ]	knave (cards)
Word-final vowel	[ẹ]	*pote*	[ˈpɔtẹ]	pot
Word-final vowel	[ẹ]	*pare*	['paɾẹ]	stop (V)
Word-final vowel	[ẹ]	*pase*	[ˈpasẹ]	pass
Word-final vowel	[ẹ]	*paxe*	[ˈpaʃẹ]	page
Word-final vowel	[o̰]	*pazo*	[ˈpaθo̰]	pazo
Word-final vowel	[o̰]	*peto*	[ˈpeto̰]	pocket

Word-final vowel	[ǫ]	*pozo*	[ˈpoθǫ]	well (N)
Word-final vowel	[ǫ]	*pucho*	[ˈpu͡tʃǫ]	calf
Word-final vowel	[ǫ]	*sapo*	['sapǫ]	toad
Word-final vowel	[ǫ]	*saco*	['sakǫ]	sack bag
Word-final vowel	[ǫ]	*sito*	['sitǫ]	situated
Word-final vowel	[ǫ]	*suco*	['sukǫ]	furrow

Experiment 1 text

Galician version: O vento e o sol

O vento do norte e mailo sol porfiaban sobre cal deles era o máis forte, cando cadrou de pasar un viaxeiro envolto nunha longa capa azul. Conviñeron en que o que antes conseguise facerlle quitar a capa ao viaxeiro sería considerado o máis forte. *Comezaron a pensar na súa mellor estratexia para gañar o reto xa que ambos eran moi competitivos e por fin decidiron que facer*[1]. O vento do norte soprou con gran furia, e canto máis sopraba máis se envolvía o viaxeiro na súa longa capa azul; finalmente o vento do norte abandonou o seu empeño. Entón o sol quentou con forza e inmediatamente o viaxeiro sacou a capa. E daquela o vento do norte tivo que recoñecer a superioridade do sol.

English version: The North Wind and the Sun

The North Wind and the Sun were disputing which was the stronger, when a traveller came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveller take his cloak off should be considered stronger than the other. *They began to think about their best strategy to win the challenge as they were both very competitive and they finally decided what to do*[1]. Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveller fold his cloak around him; and at last the North Wind gave up the attempt. Then the Sun shined out warmly, and immediately the traveller took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.

Appendix C: R commands used to fit the regression models.

Experiment 1

Mid vowels

lmer(euclid_z ~ group*style*origin + group*style*vowel + group*style*sex + style*origin*vowel + style*origin*sex + origin*vowel*sex + (1+style+vowel|participant), data=prod_mid)

Sibilant fricatives

lmer(CoG_z ~ group*phoneme*origin + group*phoneme*style + group*phoneme*sex + phoneme*origin*style + phoneme*origin*sex + origin*style*sex + (1+style+phoneme|participant) + (1+group+origin+sex|word), data=prod_fric)

Word-final vowels

lmer(f1_z ~ group*style*origin + group*style*vowel + group*style*sex + style*origin*vowel + style*origin*sex + origin*vowel*sex + (1+style+vowel|participant) + (1+group+sex+origin|word), data= prod_final)

[1] This sentence was added to increase the number of instances of key variables.