Cardiff Metropolitan University
Prifysgol Fetropolitan Caerdydd

B.Sc.(Hons) Speech and Language Therapy

Galician-Spanish mid vowel production in Southern Galicia: using language mode to explore bilingual speakers’ underlying phonological categories.

Laura López-Bueno
April 2017

Dissertation submitted in partial fulfilment of the requirements of Cardiff Metropolitan University for the degree of Bachelor of Science
DECLARATION

I hereby declare that this dissertation is the result of my own independent investigation under the supervision of my tutor.

The various sources to which I am indebted are clearly indicated. This dissertation has not been accepted in substance for any other degree and is not being submitted concurrently for any other degree.

Candidate’s signature: ________________________

RESEARCH DATA PRIVACY

I acknowledge the issue of research data privacy and undertake not to share research data in any form without the explicit approval of their supervisor.

Candidate’s signature: ________________________
Acknowledgements

Firstly, I would like to thank Dr. Robert Mayr for all the time, guidance and invaluable support he offered as my project supervisor.

I would also like to thank University College London PhD candidate Gisela Tomé-Lourido for her expert advice and guidance on Galician phonetics.

To all the Galician people who so kindly gave away their free time to collaborate in this project, my most grateful thanks for making this happen. In particular, I would like to thank those dear friends who helped locate participants for this study. Moitas grazas.

A special thank you to my brother for his ongoing support and for programming the randomiser application that was so valuable in the preparation of the experimental materials.

Finally, thanks to my mum, family, and friends for their support and presence throughout.

Thank you / Graciñas.

“Every language is a living museum, a monument to every culture it has been vehicle to.”

(Nettle & Romaine, 2000:4)
Abstract

This study analyses the production of the Galician close/open mid vowel contrasts /e-ɛ/ and /o-ɔ/, and the Spanish mid vowels /e/ and /o/ by 25 highly proficient Galician-Spanish bilinguals in the city of Vigo (Galicia, Spain) in order to investigate the impact of language switching on these vowels. While the Galician mid vowels have been studied in non-switched and code-switched paradigms, underlying phonetic interaction may have been partially obscured by connected speech effects such as speech planning and pragmatics. The present study made use of a cued picture-naming task, with bilinguals naming pictures in Galician and Spanish. Stimuli were presented in three contexts (language modes), with two separate monolingual modes –Galician only stimuli and Spanish only stimuli–, and one bilingual mode, mixing both Galician and Spanish stimuli. The tasks aimed to test whether transient effects from cross-linguistic interference would alter the phonological categorisation of the vowels when produced in the code-switched bilingual language mode as compared to the monolingual language mode. Bilinguals formed two groups according to language dominance patterns –Galician-dominant and Spanish-dominant. Results showed no significant transient effects as a consequence of language-switching. Interestingly, none of the groups made the Galician mid vowel contrasts for either the /e-ɛ/ mid vowel pair or the /o-ɔ/ mid vowel pair in either of the language mode conditions. These results have been explained by reference to a possible cross-linguistic vowel neutralisation process being on course as a consequence of long-term language contact between the two languages. Specific reference is made to variable bilingual experience as connected to multiple internal and external psycholinguistic and sociolinguistic variables, including age of acquisition, the quality and quantity of the linguistic input, habitual language use, accent, language substrate effects, and social indexicality, with a comment on the usefulness of sociophonetic models of lateral transfer to frame bilingual speech research.

This study may suit a journal with a focus on studies of an experimental nature in the general topic areas of phonetics and speech acoustics, such as the Journal of Phonetics.
Table of Contents

I. Introduction 5

II. Literature Review 5

III. Methods 11

IV. Results 20

V. Discussion 34

VI. Conclusion 40

VII. References 41

VIII. Appendices 46

i. Ethics Approval Letter 47

ii. From Methods: Complementary information
   a. Participant information sheet
   b. Consent form
   c. Participant background questionnaire
   d. Target word lists

iii. From Results: Complementary information
   a. Full data results for linear mixed models with boxplots

iv. Word count
**Introduction**

Long-term contact languages present a chance to study the phonological changes that may have developed as a consequence of cross-linguistic interaction (Mayr, Morris, Mennen and Williams, 2015, Mora et al., 2015; Simonet, 2014). The Galician and Spanish vowel systems are ideal objects of study to investigate phonological cross-linguistic phenomena (Amengual and Chamorro, 2015). Galician bilingual speakers present specific vowel production patterns which are tightly linked to sociolinguistic factors (Tomé-Lourido and Evans, 2015), so a study that reduces those variables may unveil how robustly these bilinguals realise vowel categories.

**Literature review**

*The bilingual’s two phonological systems*

Bilinguals have been shown to be highly efficient at separating the phonetic systems of their two languages, in both perception (for example, Flege, 1995) and production (see for example, Caramazza, Yeni-Komshian, Zurif and Carbone, 1973; MacLeod and Stoel-Gammon, 2010). However, the specific features of these two phonetic systems have been found to vary between speakers according to the phonological representations they have assigned to the languages’ phonetic categories (Sebastián-Gallés, Rodríguez-Fornells, Diego-Balaguer and Díaz, 2006; Amengual, 2015). The development of native-like abstract categories for these languages will depend on a wide range of variables, including the sequence in which the bilingual speakers acquire both languages (Pallier, Collomé and Sebastian-Gallés, 2001), the amount of input they are exposed to for each of the languages throughout their lives (Ellis 2002), the frequency with which they use each language (Mora, Keidel and Flege, 2010; Amengual, 2014), the features of the accent they are exposed to within their communities (Sebastián-Gallés et al., 2006; Mora, Keidel and Flege, 2015), and dialectal variation (Recarens and Espinosa, 2006; Regueira, 2009), to name but a few.

*Cross-linguistic phonological interaction and language contact*

Sequential bilinguals learn their second language (L2) after they have already started to learn their first language (L1). In sequential bilingualism, the speakers’ phonological systems have been claimed to be generally more robust for their L1 than their L2, in that speakers are often found to be more proficient at perceiving the phonetic categories of their L1 as well as producing more native-like speech (e.g. Sebastián-Gallés et al., 2006; Flege and
The study of second language speech is a growing field of research that is increasing in importance due to the contribution that non-native acquisition offers to models of human language and cognition (see Colantoni, Steele and Escudero, 2015 for a comprehensive account). A central focus of this study has been to investigate how bilinguals’ individual linguistic experiences, and particularly the age in which the L2 was acquired, may contribute to different degrees in the bilinguals’ ability to produce L2 phonology. A speakers’ L1 and L2 phonological systems do not function as independent entities, but are interrelated or combined systems in constant cross-linguistic interaction (e.g. Amengual and Chamorro, 2015; Amengual, 2012; Flege, MacKay and Meador, 1999). Age of acquisition is a central factor in this cross-linguistic interaction, and it has been studied to affect L2 learning in different ways, either because of a decrease in cerebral plasticity during normal maturation (DeKayser, 2000) or because of its relationship to other factors. Some such correlated factors are the amount of language input the speaker has been exposed to (Mora et al. 2010; Flege and Liu, 2001) and the differing amounts and frequency of use of the L1 and the L2, which partly compound the bilingual’s profile of language dominance (Birdsong, 2004; Birdsong, Gertken and Amengual, 2012). Late L2 learning onset has also been identified as a significant factor contributing to overall degree of perceived foreign accent (Piske, Mackay and Flege, 2001). Although studies tend to demonstrate that early bilinguals are more equipped to learn to discriminate and produce L1 speech to a native-like level, there is no consensus over the precise source of this potential advantage (Flege and Mackay, 2004; Flege, Yeni-Komshian and Liu, 1999). Moreover, studies have found that early experience may not be enough to acquire native-like phonetic competence (Mora et al., 2010; Pallier, Bosch and Sebastian-Gallés, 1997; Sebastián-Gallés and Bosch, 2005), and that factors such as L1-L2 dominance patterns (Amengual and Chamorro, 2015) or language-intrinsic processes such as vowel convergence (Mora et al., 2015; Best, 1995; Flege 1995, 2007) will affect the bilingual’s phonological patterns of speech in both languages.

Cross-linguistic interaction has been documented as being distinctively modulated within communities where two languages have been in long-term historical contact, and in which bilingual speakers may have been exposed, albeit variably, to both languages from birth (Mayr, Morris, Mennen and Williams, 2015, Mora et al., 2015; Simonet, 2014). Individual profiles of age of language acquisition, language use and language dominance have been found to influence bilinguals’ reception and production of both L1 and L2 phonological categories in a variety of different ways, some of which may be specific to historical factors from the relevant communities, or be embedded in a process of diachronic change as a result of long-term language contact. Two languages in long-term contact often develop structural similarities (Bullock and Gerfen, 2004; Chang, 2009), and phonetic convergence is a process well documented consequence of language contact (for example Mayr et al., 2015 for
the Welsh-English pair; Bullock and Gerfen, 2004 for the French-English pair; Recarens and Espinosa, 2009 for the Catalan-Spanish pair). A wealth of literature has been dedicated in recent times to explore the specific phonological behaviours of long-term contact language communities. The Catalan-Spanish contact language pair has been a particularly prolific focus of research in recent times (Amengual, 2016; Mora et al., 2015, 2010; Recarens and Espinosa, 2009; Simonet, 2014, etc) and findings are being explored for other peninsular language pairs sharing similar phonetic system structures, such as the Galician-Spanish phonetic systems (for example, Amengual and Chamorro 2015, Tomé-Lourido and Evans 2015).

The sociophonetics of L2 Galician speech

The Spanish language is an increasingly dominant language in society respective to Galician, particularly in larger urban centres such as Vigo (RAG, 2004; Rei-Doval, Fernández-Rodríguez, Rodriguez-Neira et al. 1994). The complex sociolinguistic variables surrounding Spanish-Galician bilingualism have been extensively researched and described in the literature corpus of Galician studies. This complexity stems in part from the co-existence of Galician and Spanish as historical contact languages, where Galician was placed in a situation of strong diglossia as the low prestige language in respect to the higher-prestige Spanish variety (Ferguson, 1959) for centuries on end, which intensified during Franco’s dictatorship in the mid decades of the 20th century. Active political action targeted and eliminated the use of Galician in academia, the institutions and schooling, which caused a generational gap in bilingual language acquisition. Urban generations who were less in contact with the spoken varieties of Galician were raised and educated as Spanish monolinguals. Spoken Galician was relegated to the lower classes and people living outside the urban centres, who were labelled uneducated and as a consequence increasingly spoke Spanish to their children as an attempt to protect them from future social inequalities caused by the use of Galician. Since the late 1970s, new cultural and political efforts to protect the Galician language, which was seen to rapidly lose its older speaker base and at risk of becoming extinct, brought about policies that enacted compulsory primary and secondary schooling in the Galician language and promoted its use and transmission. This was efficient in bringing about a new generation of L2-Galician bilingual speakers (see Turell, 2001 for a comprehensive historical account of Galician sociolinguistics).

Over the following three decades, a new social group of speakers has been identified, Neofalantes (‘new speakers’) in the literature (O’Rourke and Ramallo 2015, 2013), although their precise classification criteria are a matter for debate. These are early Spanish L1 bilinguals who have chosen to adopt the use of the L2 Galician language that
they had learned in school but rarely used, to differing levels of competence and confidence, out of a strong ideology and sense of identity with the minority language and the Galician cultural heritage. In changing their habitual use of Spanish to speak Galician as their new language of choice, they have made a conscious effort to change language dominance, and often express motivation to use Galician exclusively (O’Rourke and Ramallo, 2015). Sociolinguistic findings since Labov (1966) have shown that bilingual speakers are aware of phonological forms in relation to how these are associated with particular social groups and situations, and are able to vary their production to signal attributes such as status and prestige, being this one of the main vehicles of diachronic language change (Labov, 1972). In the case of the Galician sociolinguistic situation, this raises the question of whether there is a relationship between the speaker’s chosen indexicality as a member of a social group, and how faithfully they produce the categorical phonetic elements of both their languages, as well as what the effects of language dominance are on their phonological behaviours.

Tomé-Lourido and Evans (2015) investigated the effects on the perception and production of the Galician mid-vowel contrast (/e- ɛ/ and /o- ɔ/) of a change in language dominance for ideological reasons by three groups of Galician-Spanish bilinguals living in Santiago de Compostela, in Northern Galicia, with particular focus on the Neofalantes group. Results showed that these speakers produced intermediate categories which were neither the Spanish nor the Galician categories. They hypothesized that this could point to an attempt to use indexicality through phonological means, as a way of distinguishing themselves from either of the other groups and adopt a position in between both, retaining features from their previously dominant language. The authors provide a second possible explanation of the results as potentially being a consequence of L2 learning effects rather than dominance, in that new speakers have not fully acquired the phonetic categories due to late age of acquisition and reduced flexibility in both perception and production of the L2 categories. However, they also point to a wide range of individual variation in Galician mid-vowel production. Galician dominant speakers were found to be able to produce both mid-vowel contrasts distinctively, while Spanish dominant speakers followed the opposite pattern and were unable to produce these mid-vowels contrastively as phonetic categories, producing a merged category for each of the vowel pairs.

Language dominance effects were also studied by Amengual and Chamorro (2015) for the same mid-vowel pairs. They tested the perception and production of the Galician /e- ɛ/ and /o- ɔ/ mid-vowel contrast in two groups of early Galician and Spanish bilingual speakers living in the cities of Santiago de Compostela and Vigo. They looked at speakers’ mid-vowel categories in relation to their individual profiles of language dominance, and interestingly found that Spanish dominant speakers had merged front mid vowels, but were however able to
produce separate categories for the back mid vowels. The authors comment further on the nature of language dominance in bilingual speakers, and how both indexical factors of identification with the speakers of a language and the amount of use of this language may be defining elements in relation to their ability to produce the vowel pair categories.

These seemingly contradictory findings partly illustrate the complexity of disentangling sociolinguistic factors from purely linguistic ones in the study of bilingual speech. The use of experimental design to control elements of the tightly linked variables in order to explore how the linguistic phenomena responds in the presence or absence of those elements may greatly contribute to a better understanding of how these variables influence or interfere with the speakers’ phonetic and phonological behaviours, in turn demonstrating how these are ultimately inseparable from sociolinguistic factors under more naturalistic conditions (Thomas, 2011).

One such variable that could be considered in the light of previous experimental design on bilingual phonology is the communicative context in which the language is being used. Research shows that, while perception may not be very much affected by language context (Caramazza et al., 1973; Flege and Eefting, 1987), the effects of language context on production are considerable (Antoniou, Best, Tyler and Kroos, 2010; Magloire and Green, 1999). Would this be a factor influencing Galician-Spanish bilinguals’ production of the mid-vowel contrasts, and would it be possible to isolate the contextual variables in a single experiment in order to explore phonological behaviour without contextual constraints?

Language modes and language switching

Grosjean’s Bilinguals’ Language Modes framework (1982, 1989, 1998b, 2001, 2008, 2012) proposes that, in regards to language processing mechanisms, a bilingual’s two languages are unequally activated at any one time, and the amount of activation each of them has depends on the communicative setting in which the bilingual is engaged. How strongly the languages are activated may affect speech behaviour. When the speaker is in a monolingual mode, either one of the languages will be more strongly activated, and phonetic behaviour will be more robust (Grosjean and Li, 2012). In a bilingual mode, when there are contextual effects for both languages at once, competing for the bilingual’s cognitive attention mechanism, cross-linguistic interference effects may arise, being these either static, i.e. long term, such as a foreign accent, or dynamic, i.e. short term, such as the phonology of the L1 interfering with the L2 during code-switching (Grosjean, 2012). Transient activation effects have not been consistently found in studies (Antoniou et al., 2011; Grosjean & Miller, 1994; Olson, 2013). However, long-
term a more robust longer lasting effect of activation has been found in bilinguals’ speech when language mode activation is prolonged (Flege, 1987; Olson, 2013, 2016; Sancier & Fowler, 1997). Language mode can be induced in an experimental setting by controlling for the amount of language context offered in a discourse so that the speaker is exposed to either mostly monolingual language context (however, see Grosjean, 2008, for cognitive constraints in inducing a purely monolingual mode) or a bilingual language context where both languages coincide.

A cued picture-naming experimental paradigm in which language-switching tasks and bilingual language mode are contrasted with the monolingual mode and naming may offer a good opportunity to explore transient activation effects between the Galician-Spanish bilingual speaker’s two phonological systems. Given previous findings, would the Spanish dominant production of the mid-vowel categories be replicated once context effects are controlled for and reduced? Would Galician dominant speakers’ production be affected by cross-linguistic context effects, and would their mid-vowel contrast production be as robust as demonstrated in previous studies? Would language mode produce any phonological effect on how speakers realise phonetic categories in either language as a consequence of language switching? The purpose of the present study was to address these questions.

Methods

Introduction

The present study investigated the production of the Galician mid vowel contrasts (/ɛ̃-ɛ̃/ and /ɔ̃-ɔ̃/) and the Spanish mid vowels (/e/, /o/) in monolingual and code-switched language mode conditions by 25 Galician-Spanish bilinguals living in the city of Vigo in Galicia. The main goal was to investigate the potential effects of language switching and language mode on bilingual phonetic production, specifically in the context of historical contact languages.

Participants

Recruitment

A total of 38 Galician-Spanish bilinguals participated in a cued picture-naming task followed by a questionnaire to assess their language dominance and linguistic experiences. Potential participants were initially contacted using
a combination of direct contact and snow-balling via personal phone calls, email messages and social media in the researcher’s personal circles. The recruitment process continued on location throughout the data collection period via word of mouth and the researcher’s direct contact with local cultural organisations. Recruitment was discontinued at the end of the three-week period allocated to data collection.

Initial Selection and criteria

All participants were adults, male and female, with ages ranging between 26 and 44 years old (Female n=24, age range=26-44; Male n=14, age range=37-43). Participants were all bilingual Galician-Spanish speakers with a variety of linguistic backgrounds and patterns of language use. None were native speakers of other languages. All were residents in the city of Vigo (Southwestern Galicia) and surrounding urban areas.

Characteristic usage patterns in the Galician phonology of speakers living in the western Galician regions have been found to differ from those of speakers living in the eastern Galician regions (Costas Gonzalez, 1988; Regueira, 2008). Relevant to this study are the specific differences found in the production of mid-vowel contrasts, where a word that is typically pronounced with an open mid-vowel in eastern regions may be pronounced by some speakers with a close mid-vowel in western regions due to a phonological process of ‘nominal metaphony’ (Regueira, 2008). Due to potential irregularities in pronunciation patterns between the two areas, only participants that had lived in the western areas of Galicia for a substantial part of their lives were selected for this study.

All participants attended compulsory schooling in western Galicia from age 4 or 5 to at least 14 years of age. Most completed higher education. At the time when the participants were in primary education, the Galician government had officially decreed that the Galician language be taught in schools (“Decree on Bilingualism 1st of August 1979”). Due to the presence of Spanish as an increasingly dominant language in society respective to Galician in the urban centres of southwestern Galicia (RAG, 2004; Rei-Doval, Fernandez-Rodriguez, Rodriguez-Neira et al., 1994, p.58), it has been assumed no purely monolingual Galician speaker would likely exist from the age range selected. All participants were bilinguals who had learnt the Galician language either through early experience as their L1, or through late experience as an L2 during schooling.

Bilingual proficiency was determined both during the recruitment process and during data collection sessions. In order to discard potential Spanish monolinguals, Galician was used as the medium through which pre-interviews and arrangements to attend sessions were carried out. Participants were engaged in conversation using both the
Spanish and Galician languages with the researcher (a late Galician-Spanish bilingual) before and in between the experiment’s tasks. This was used to determine fluency in both languages. Participants reported different levels of confidence in their abilities to speak the Galician language fluently, but none showed any difficulty in reading written materials in Galician, all responded to the questionnaire in correct Galician writing, and all were found to present similar levels of fluency in conversation in both languages regardless of dominance.

**Language use and dominance**

Two groups of participants were distinguished: (1) Galician-Spanish bilinguals with a Galician-dominant pattern of language use, i.e. GAL DOM; and (2) Galician-Spanish bilinguals with a Spanish-dominant pattern of language use, i.e. SPA DOM. GAL DOM participants had early experience of Galician with at least one Galician-speaking parent and/or having been raised in a Galician-speaking household. SPA DOM participants had no reported early experience of Galician in the household, and acquired Galician at a later age via schooling.

The classification of participants as Spanish-dominant or Galician-dominant was determined from their responses to questionnaire questions about language choice and patterns of use. Participants who reported having Galician as their main language were pre-assigned to the GAL DOM group, while those who reported having Spanish as their main language were pre-assigned to the SPA DOM group. Language dominance was confirmed from self-reported amount of Galician use in 16 possible domains (social environments and relationships). Participant answers differed in the number of domains of use that were relevant to them at the time of the study. Equal score weighting was given to each of these domains in a five-point Likert-style scale reporting level of use ranging from “Never” to “Always”, following Gertken, Amengual and Birdsong’s (2014) scoring method for assessing bilingual language dominance.

A percentage of use (PU) was calculated from the ratio formed by the total number of domains for each level of use (U) and the total number of domains reported (TD), i.e. \( PU = \frac{U}{TD} \)
Table I depicts the result percentages of use of Galician according to participant’s reported use across domains and given reported main language. Inspection of the table shows that the two groups clearly differ in the amount of Galician use reported as “Always” and “Often”, with participants who report Galician being their main language reporting the largest percentages, and those who identify as being mainly speakers of Spanish reporting the lowest percentages. Galician language dominance was then defined as a combined percentage of more than 50% use for the “Always” and “Often”. Spanish language dominance was defined as a percentage of use of Galician of less than 50% in the same combined categories.

**Final selection**

A total of 12 participants were discarded for not fitting the criteria specified above. One participant was discarded due to a data collection error, so these data have not been included in the account below.

Data were analysed from a final total number of 25 participants (Female=16, Male=9). Two groups were formed with Galician-dominant, early Galician experience bilinguals (GAL DOM, n=15; Female=10, Male=5) and Spanish-dominant, no early Galician experience bilinguals (SPA DOM, n=10; Female=4, Male=6).

None of the participants reported medical conditions or sensory difficulties that could have interfered with the experimental process.
Data Collection

Target words and experimental design

48 Galician and Spanish words (Galician=32, Spanish=16) were selected containing the target Galician mid-vowel pairs /e- ɛ/ and /o- ɔ/ and the target Spanish mid vowels /e/ and /o/, resulting in 8 target words per vowel (ILG USC, 2016 for Galician pronunciation). All target vowels were embedded in the tonic first syllable of a bisyllabic word with a prosodic environment of ‘(C)VC(C)V). Words were matched in syllabic structure and general prosodic pattern as much as possible in order to reduce phonetic context effects. Due to their historical relationship as languages in contact, the Galician and Spanish vocabularies present with a high number of cognates where the only phonetic change between the pairs is a difference in the mid-vowel. In order to minimise language attribution errors in the language switching task, Spanish target words were selected to contain at least one other phonetic difference besides the vowel with respect to their Galician translation. Target words were all nouns, of high frequency where possible to facilitate naming (Ellis, 2002) (for Galician frequency CRP, 2015; for Spanish frequency RAE, 2016). Further to this, most targets were chosen to have good imageability to facilitate naming. To assess naming consistency and identify potential problems, five non-participant volunteers were asked to name all images of target words. Changes to images misnamed or reported as ambiguous were incorporated in designing the final stimuli sets.

Target words were named by participants in three ‘language mode trials’ (lists of targets + fillers in different combinations), i.e. (1) a Spanish Monolingual mode trial; (2) a Galician Monolingual mode; and (3) a Bilingual mode. Trials (1) and (2) contained only targets and fillers for one of the languages while (3) combined targets and a reduced number of fillers from both languages.

Participants (n=25) named 76 target words each, producing a total of 1900 target productions. Each participant named each target word twice over the three language mode trials: once in a monolingual mode trial, once in a bilingual mode trial.

A reduced set of target words for each of the target vowels was selected for full analysis. Table III depicts the selected list of target words that was analysed for this study. Since participants made a large amount of naming errors, target words were selected so that the most consistently named were also the ones chosen to be analysed. One vowel set has 6 target words so that the total number of tokens analysed is similar to the other sets.
Table II. Target tokens in Galician (GAL) and Spanish (SPA) containing the target vowels in the first syllable position. English translation provided.

<table>
<thead>
<tr>
<th>GAL /ɔ/</th>
<th>GAL /o/</th>
<th>GAL /ɛ/</th>
<th>GAL /e/</th>
<th>SPA /o/</th>
<th>SPA /e/</th>
</tr>
</thead>
<tbody>
<tr>
<td>roda</td>
<td>goma</td>
<td>sete</td>
<td>cebra</td>
<td>ropa</td>
<td>beso</td>
</tr>
<tr>
<td>porto</td>
<td>mono</td>
<td>serpe</td>
<td>cebra</td>
<td>hombro</td>
<td>pero</td>
</tr>
<tr>
<td>foca</td>
<td>goma</td>
<td>regra</td>
<td>cepo</td>
<td>foco</td>
<td>beso</td>
</tr>
<tr>
<td>nove</td>
<td>bone</td>
<td>tecla</td>
<td>seto</td>
<td>fonte</td>
<td>beso</td>
</tr>
<tr>
<td>óso</td>
<td>oso</td>
<td>césar</td>
<td>testo</td>
<td>caesar</td>
<td>ceja</td>
</tr>
<tr>
<td>peña</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>teja</td>
</tr>
</tbody>
</table>

Stimuli

Target stimuli for the picture-naming study consisted of 48 target pictures and 40 filler pictures, in black and white line drawing style. Most pictures were taken from Snodgrass and Vanderwart (1980), and the rest from a variety of similar picture resources. Target images were selected taking into account non-ambiguity of interpretation to facilitate a high level of consistency in target word naming by all participants.

Procedure

Data were collected in Vigo (Galicia) over a period of 3.5 weeks. Each participant attended a one-hour session at a quiet designated place of their choice, in private premises. Each session consisted of an introduction and consent-taking from the participant, time to respond to queries, the experimental task(s), and the language background questionnaire, in this order. Participants were informed that the study involved naming words from pictures in order to study how bilingual speakers use both languages when they speak. All consented to their speech being recorded for later acoustic analysis.

Each participant produced 3 recordings, one of each language mode trial (2 x monolingual mode trials, and 1 x bilingual mode trial) (See above). Voice tracks were recorded in WAV format using a Zoom H2n Handy Recorder with integrated microphone.

Participants were greeted to the session in the Galician language. This was intended to informally determine the individual participant’s level of proficiency and fluency in this language. Instructions were given verbally in the language coinciding with the language mode trial they were about to complete. This in turn was intended to influence a language mode (monolingual or bilingual) in the participant in line with the trial at hand (Grosjean,
e.g. 2001). Before the bilingual trial, verbal code-switching behaviour was used by the researcher for a period of time with the participants. Assignment of trial order was counter-balanced across individuals to ensure there were no order effects on the data. Token words (targets + fillers) within each trial were also randomised ahead of the recording period using a self-designed randomiser macro in Powerpoint (Microsoft, 2016), so no two trials were the same for any of the participants. Neither researcher nor participant could predict the order in which the tokens would appear.

Participants completed a familiarisation stage before the experimental stage. This was done in the language of the trial, and consisted of three smaller sets of discarded token words (one per language mode trial), which were also randomised. Participants were asked to name the pictures as accurately as possible and with a natural voice. Stimuli were displayed on a computer screen at a distance that was reported to be comfortable by all participants. Image I below shows an example of two token stimuli used for this experiment.

Participants were instructed to name the tokens in the language indicated next to the token picture. The language to be selected was specified by the presence of one of two cartoon men, one dressed in blue (Galician flag colour) and representing the Galician language, and the other dressed in light red representing the Spanish language (see Image I). This design was chosen to offer the maximum amount of visual cues that the participant could associate to the language of reference, in order to aid the memory retention of the sign during the production tasks. It was expected this would also minimise interference of other cognitive processes in the language selection mechanism, e.g. participants engaging in attempts to retrieve the sign’s referent from memory instead of devoting attentional processes to language selection.
Analysis

Acoustic analysis

A total of 31 target words per participant were analysed, over 2 language mode trials (total=62), multiplied by the number of participants (n=25), led to 1550 target words attempted. 100 of the targets were either not produced by the participants or misnamed, leaving a grand total of 1450 productions for analysis: GAL /o-ɔ/=493 productions (/o/=244 productions, /ɔ/=249 productions); GAL /e-ɛ/=476 productions (/e/=211 productions, /ɛ/=264 productions); SPA /o/=238 productions; SPA /ɛ/=243 productions.

The words were analysed acoustically using PRAAT software 5.1.04 (Boersma and Weenink, 2017). Target vowels were isolated by hand from neighbouring segments, from the beginning of the first well-formed period of the vowel to the end of the last well-formed period. Boundaries were adjusted to the zero crossing automatically using the dedicated PRAAT function. Whenever boundaries were unclear, e.g. when the vowel was placed between two nasal consonants, the boundaries were decided by visually inspecting the movement of the second formant (F2), the waveform patterns, and by aurally marking the difference between the segments. All coding was performed by the researcher to eliminate inter-coder variation.

Image II. Screenshot of PRAAT software showing vowel segmentation of Galician /ɛ/ in the word “serpe” (snake)
The formant tracking function in PRAAT was used to measure the first and second formants (F1, F2) of each vowel at their mid-point, located using the programme’s automatic function. All vowels were monophthongs with negligible vowel-inherent spectral change. Settings were adjusted to a ceiling of 5000 Hz for male voices and 5500 Hz for female voices in order to adjust for the average differences in male and female vocal tract sizes which result in higher formant values for women (Colantoni, Steele and Escudero, 2015), with a dynamic range of 35dB and a 5 formant tracking selection. All outlier F1 and F2 automatic measurements were hand corrected when they showed mistracking. Raw Hertz values were converted into Bark (Traunmüller, 1990) in order to allow for comparisons of data from speakers, cancelling the effects of physiological differences.

Preparation for Statistical Analysis

All relevant data, vowel formant measurements and their converted values were coded and inputted in 4 separate documents, two for the Galician vowel pairs /e-ɛ/ and /o-ɔ/, and two for the Spanish vowels /e/ and /o/, ready to be explored using the statistical package R (see Results section for full description).

Ethical considerations

All participants were presented with a participant information sheet at the beginning of the session. They all signed a participant consent form, and were reminded verbally of their right to withdraw from the study. A copy of both documents is included in the appendix section.
Results

The present study investigated the production of the Galician mid vowel close/open contrasts (/e-ɛ/ and /o-ɔ/) and the Spanish mid vowels (/e/, /o/) in monolingual (Galician, Spanish) and code-switched (bilingual) language mode conditions, by 25 Galician-Spanish bilinguals participating in a cued picture-naming task. Differences between vowel productions were studied via the first and second formant Bark-converted values (B1, B2) at each target vowel’s midpoint. Bark-converted formant data were coded to reflect the dominant language group of the participant who produced each value (Group= Galician dominant; Spanish dominant), the language mode in which it was produced (Mode= Monolingual mode [Galician/Spanish]; Bilingual mode), and the vowel type when appropriate (Vowel= each type of the vowel pair [as below]; no vowel code).

To determine differences between the language dominance groups, language modes and target vowels, linear mixed effects models were run separately in R (R Core Team, 2016) for each of four groups of acoustic data: (1) the Galician front close/open mid vowel pair /e-ɛ/; (2) the Galician back close/open mid vowel pair /o-ɔ/; (3) the Spanish front mid vowel /e/; (4) the Spanish back mid vowel /o/.

For models (1) and (2), group, mode and vowel were entered as fixed factors (including interaction) and Participant as a random factor with random slopes for Vowel. For models (3) and (4), Group and Mode were entered as fixed factors (including interaction) and Participant as a random factor.

P-values were generated via degrees of freedom obtained through the Satterthwaite approximation, using the LmerTest function in R (Bates, Maechler, Bolker, & Walker, 2015).

First formant (F1) vowel analysis, corresponding to vowel height, is the focus of this study and hence it will be reported first within each section. Complementary analysis of the vowels’ second formant (F2) will also be reported.

Results of the analysis of the Galician vowel pairs will be described in the first place, followed by results for the Spanish single vowel analyses.

The appendix section contains additional boxplots and data results from the models in full.
1. **Galician front close/open mid vowel pair /e-ɛ/**

*F1 (B1).* Figure 1 shows the boxplot representing the F1 production of the Galician vowel pair /e-ɛ/ by language dominance and language mode. Both groups presented a wide range of F1 frequencies, the median frequency being very similar across group, vowel and language mode. The group that presented the largest variation in first formant frequency value range was the GAL DOM group for the Galician front open mid vowel /ɛ/.

![Figure 1. First formant (F1) Bark-converted distributions for Galician /e/ and /ɛ/ by participants' language dominance (GAL DOM = Galician dominance; SPA DOM = Spanish dominance) and language mode (Monolingual Galician mode = White; Bilingual mode = Blue).](image)

To determine whether there were significant differences in the vowel height dimension when the targets were produced by each language dominance group in the separate language mode conditions, a linear mixed-effects model was run for the F1 data, with Group (2 levels: GAL DOM; SPA DOM), Mode (2 levels: Monolingual Galician Mode; Bilingual Galician-Spanish Mode) and Vowel (2 levels: /e/; /ɛ/) as fixed factors (including interaction) and Participant as a random factor with random slopes for Vowel. The results, displayed in Table 1, showed no significant main effect of Group, Mode or Vowel. They also showed no significant Group*Vowel, Group*Mode, Mode*Vowel or Group*Mode*Vowel interactions. The model results showing no significant vowel effect suggest neither group made the vowel distinction. The Spanish dominant group produced more open vowels for both types and in both modes when compared to the Galician dominant group, who produced more closed mid vowels across the board.
Table 1. Results of mixed-effects model for Galician vowel pair /e-ɛ/ Bark-converted first formant (F1)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>df</th>
<th>t</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>4.62510</td>
<td>0.13888</td>
<td>35.5000</td>
<td>33.302</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>Group</td>
<td>0.27344</td>
<td>0.17961</td>
<td>35.7000</td>
<td>1.522</td>
<td>0.137</td>
</tr>
<tr>
<td>Mode</td>
<td>-0.03582</td>
<td>0.11369</td>
<td>450.1000</td>
<td>-0.315</td>
<td>0.753</td>
</tr>
<tr>
<td>Vowel</td>
<td>0.10840</td>
<td>0.10734</td>
<td>439.1000</td>
<td>1.010</td>
<td>0.313</td>
</tr>
<tr>
<td>Group*Vowel</td>
<td>0.04629</td>
<td>0.13962</td>
<td>439.7000</td>
<td>0.332</td>
<td>0.740</td>
</tr>
<tr>
<td>Group*Mode</td>
<td>0.13655</td>
<td>0.14707</td>
<td>450.1000</td>
<td>0.928</td>
<td>0.354</td>
</tr>
<tr>
<td>Mode*Vowel</td>
<td>0.11942</td>
<td>0.15197</td>
<td>450.1000</td>
<td>0.786</td>
<td>0.432</td>
</tr>
<tr>
<td>Group<em>Mode</em>Vowel</td>
<td>-0.13883</td>
<td>0.19708</td>
<td>450.1000</td>
<td>-0.704</td>
<td>0.482</td>
</tr>
</tbody>
</table>

To investigate whether there were effects of by-item variation, and hence discard errors of the application of the independence assumption (Winter, 2013), a further linear mixed effects model was run following the model as above, with an additional random slope for Item with Participants as random factor (1+Item|Participant). Results again showed no significant main effect of Group (t(38.5) = 1.52, p = .137) Mode (t(450) = -0.31, p = .753), Vowel (t(429) = 1.01, p = .311), or any of the corresponding interactions. This suggests both groups produced no significant item-specific vowel differences by either vowel type or language mode condition.

F2 (B2). Figure 2 shows the boxplot representing the F2 production of the Galician vowel pair /e-ɛ/ by language dominance and language mode. Both groups produced a very similar median frequency across group, vowel and language mode, with a rather large variability in the range of productions overall, in particular the GAL DOM group.
To determine whether there were significant differences in the vowel frontness dimension when the targets were produced by each language dominance group in the separate language mode conditions, a linear mixed-effects model was run for the F2 data, with Group (2 levels: GAL DOM; SPA DOM), Mode (2 levels: Monolingual Galician Mode; Bilingual Galician-Spanish Mode) and Vowel (2 levels: /e/; /ɛ/) as fixed factors (including interaction) and Participant as a random factor with random slopes for Vowel. There were no significant main effect of Group, Mode or Vowel. The results also showed no significant interactions of Group*Vowel, Group*Mode, Mode*Vowel or Group*Mode*Vowel.

Table 2. Results of mixed-effects model for Galician vowel pair /e-ɛ/ Bark-converted second formant (F2)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>df</th>
<th>t</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>12.74764</td>
<td>0.21413</td>
<td>30.7000</td>
<td>59.532</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>Group</td>
<td>0.39154</td>
<td>0.27675</td>
<td>30.8000</td>
<td>1.415</td>
<td>0.167</td>
</tr>
<tr>
<td>Mode</td>
<td>-0.17867</td>
<td>0.13726</td>
<td>425.1000</td>
<td>-1.302</td>
<td>0.193</td>
</tr>
<tr>
<td>Vowel</td>
<td>0.24425</td>
<td>0.14071</td>
<td>64.6000</td>
<td>-1.736</td>
<td>0.087</td>
</tr>
<tr>
<td>Group*Vowel</td>
<td>0.03299</td>
<td>0.18284</td>
<td>65.8000</td>
<td>0.180</td>
<td>0.857</td>
</tr>
<tr>
<td>Group*Mode</td>
<td>0.30425</td>
<td>0.17753</td>
<td>424.4000</td>
<td>1.714</td>
<td>0.087</td>
</tr>
<tr>
<td>Mode*Vowel</td>
<td>0.22365</td>
<td>0.18346</td>
<td>425.9000</td>
<td>1.219</td>
<td>0.223</td>
</tr>
<tr>
<td>Group<em>Mode</em>Vowel</td>
<td>-0.46229</td>
<td>0.23788</td>
<td>425.0000</td>
<td>-1.943</td>
<td>0.053</td>
</tr>
</tbody>
</table>
Within-item differences were investigated via a linear mixed effects model as above that included an additional random slope for Item with Participant as random factor (1+Item|Participant). Results again showed no significant main effect of Group (t(30.8) = 1.41, p = .167) Mode (t(428) = -1.32, p = .187), Vowel (t(88.2) = -1.77, p = .081), or any of the corresponding interactions.

2. Galician back close/open mid vowel pair /o-ɔ/

F1 (B1). Figure 3 shows the boxplot representing the F1 production of the Galician back mid vowel pair /o-ɔ/ by language dominance and language mode. Both groups presented a narrower range of F1 frequencies when compared to the front mid vowel pair, with median frequencies being very similar across group, vowel and language mode. This lack of variation shows both groups produced a merged single vowel category in both language modes consistently, regardless of target vowel type.

![Figure 3. First formant (F1) Bark-converted distributions for Galician /æ/ and /β/ by participants’ language dominance (GAL DOM = Galician dominance; SPA DOM = Spanish dominance) and language mode (Monolingual Galician mode = White; Bilingual mode = Blue).](image)
To determine whether there were significant differences in the vowel height dimension when the targets were produced by each language dominance group in the separate language mode conditions, a linear mixed-effects model was run for the F1 data, with Group (2 levels: GAL DOM; SPA DOM), Mode (2 levels: Monolingual Galician Mode; Bilingual Galician-Spanish Mode) and Vowel (2 levels: /o/; /ɔ/) as fixed factors (including interaction) and Participant as a random factor with random slopes for Vowel. The results, displayed in Table 3, showed no significant main effect of Group, Mode or Vowel. In addition, they showed no significant Group*Vowel, Group*Mode, Mode*Vowel or Group*Mode*Vowel interactions. Differences in vowel categorisation had been found to be increasingly marked for the back vowel pair in Galician Dominant speakers (Amengual and Chamorro, 2015; Tomé-Lourido and Evans, 2015), with speakers whose dominant language was Spanish again producing a merged category for both the back open and close mid vowels (Gisela). Given that the current model showed no significant effects of vowel, results seemed to suggest neither group in this study produced distinctive vowel categories.

Table 3. Results of mixed-effects model for Galician vowel pair /o-ɔ/ Bark-converted first formant (F1)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>4.77361</td>
<td>0.15110</td>
<td>35.00000</td>
<td>31.592</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>Group</td>
<td>-0.07445</td>
<td>0.19516</td>
<td>35.10000</td>
<td>-0.381</td>
<td>0.705</td>
</tr>
<tr>
<td>Mode</td>
<td>0.05268</td>
<td>0.12121</td>
<td>443.30000</td>
<td>0.435</td>
<td>0.664</td>
</tr>
<tr>
<td>Vowel</td>
<td>0.10855</td>
<td>0.13422</td>
<td>67.40000</td>
<td>0.809</td>
<td>0.422</td>
</tr>
<tr>
<td>Group*Vowel</td>
<td>0.21678</td>
<td>0.17358</td>
<td>67.80000</td>
<td>1.249</td>
<td>0.216</td>
</tr>
<tr>
<td>Group*Mode</td>
<td>-0.07172</td>
<td>0.15668</td>
<td>443.20000</td>
<td>-0.458</td>
<td>0.647</td>
</tr>
<tr>
<td>Mode*Vowel</td>
<td>-0.01777</td>
<td>0.17050</td>
<td>442.90000</td>
<td>-0.104</td>
<td>0.917</td>
</tr>
<tr>
<td>Group<em>Mode</em>Vowel</td>
<td>0.18233</td>
<td>0.22041</td>
<td>442.90000</td>
<td>0.827</td>
<td>0.409</td>
</tr>
</tbody>
</table>

To investigate whether there were effects of by-item variation a further linear mixed effects model was run following the model as above, with an additional random slope for Item with Participant as random factor (1+Item|Participant). Results showed no significant main effect of Group (t(41.6) = -0.23, p = .820), Mode (t(443) = 0.44, p = .657), Vowel (t(84.2) = 1.04, p = .300), or any of the corresponding interactions. This suggests both groups produced no significant item-specific vowel differences by either vowel type or language mode condition.
Figure 4 shows the boxplot representing the F2 production of the Galician back mid vowel pair /o-ɔ/ by language dominance and language mode. Both groups presented a narrower range of F1 frequencies when compared to the front mid vowel pair, with median frequencies being very similar across group, vowel and language mode. This lack of variation shows both groups produced a merged single vowel category in both language modes consistently, regardless of target vowel type.

To determine whether there were significant differences in the vowel frontness dimension when the targets were produced by each language dominance group in the separate language mode conditions, a linear mixed-effects model was run for the F2 data, with Group (2 levels: GAL DOM; SPA DOM), Mode (2 levels: Monolingual Galician Mode; Bilingual Galician-Spanish Mode) and Vowel (2 levels: /o/; /ɔ/) as fixed factors (including interaction) and Participant as a random factor with random slopes for Vowel. There were no significant main effect of Group (t(51.3) = 0.14, p = .888), Mode (t(468) = 1.56, p = .120) or Vowel (t(176) = 1.20, p = .231). The results also showed no significant interactions of Group*Vowel (t(176) = 1.20, p = .231), Group*Mode (t(468) = -1.40, p = .162), Mode*Vowel (t(468) = -1.46, p = .145) or Group*Mode*Vowel (t(468) = 1.18, p = .239).
Table 4. Results of mixed-effects model for Galician vowel pair /o-ɔ/ Bark-converted second formant (F2)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>8.60779</td>
<td>0.11047</td>
<td>51.20000</td>
<td>77.920</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>Group</td>
<td>0.02025</td>
<td>0.14273</td>
<td>51.30000</td>
<td>0.142</td>
<td>0.888</td>
</tr>
<tr>
<td>Mode</td>
<td>0.18546</td>
<td>0.11920</td>
<td>468.00000</td>
<td>1.555</td>
<td>0.120</td>
</tr>
<tr>
<td>Vowel</td>
<td>0.09546</td>
<td>0.12369</td>
<td>175.60000</td>
<td>0.772</td>
<td>0.441</td>
</tr>
<tr>
<td>Group*Vowel</td>
<td>0.19210</td>
<td>0.15999</td>
<td>176.50000</td>
<td>1.201</td>
<td>0.231</td>
</tr>
<tr>
<td>Group*Mode</td>
<td>-0.21560</td>
<td>0.15409</td>
<td>468.00000</td>
<td>-1.399</td>
<td>0.162</td>
</tr>
<tr>
<td>Mode*Vowel</td>
<td>-0.24493</td>
<td>0.16771</td>
<td>468.00000</td>
<td>-1.460</td>
<td>0.145</td>
</tr>
<tr>
<td>Group<em>Mode</em>Vowel</td>
<td>0.25544</td>
<td>0.21681</td>
<td>468.00000</td>
<td>1.178</td>
<td>0.239</td>
</tr>
</tbody>
</table>

Within-item differences were investigated via a linear mixed effects model as above that included an additional random slope for Item with Participants as random factor (1+Item|Participant). Results again showed no significant main effect of Group (t(48.6) = -0.15, p = .878), Mode (t(445) = 1.57, p = .119), Vowel (t(97.7) = 0.78, p = .437), or any of the corresponding interactions.

3. Spanish front mid vowel /e/

F1 (B1). Figure 5 shows the boxplot representing the F1 production of the Spanish front mid vowel /e/ by language dominance and language mode. Both groups presented a variable range of F1 frequencies, the median frequency being very similar across group and language mode.
To determine whether there were significant differences in the vowel height dimension when the target vowel was produced by each language dominance group in the separate language mode conditions, a linear mixed-effects model was run for the F1 data, with Group (2 levels: GAL DOM; SPA DOM) and Mode (2 levels: Monolingual Spanish Mode; Bilingual Galician-Spanish Mode) as fixed factors (including interaction) and Participant as a random factor. The results predictably showed no significant main effect of Group or Mode, and no Group*Mode interactions.

Table 5. Results of mixed-effects model for Spanish vowel pair /e/ Bark-converted first formant (F1)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>4.65458</td>
<td>0.12865</td>
<td>34.29000</td>
<td>36.180</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>Group</td>
<td>0.29063</td>
<td>0.16655</td>
<td>34.66000</td>
<td>1.745</td>
<td>0.090</td>
</tr>
<tr>
<td>Mode</td>
<td>0.09794</td>
<td>0.10355</td>
<td>217.97000</td>
<td>0.946</td>
<td>0.345</td>
</tr>
<tr>
<td>Group*Mode</td>
<td>-0.05446</td>
<td>0.13355</td>
<td>217.93000</td>
<td>-0.408</td>
<td>0.683</td>
</tr>
</tbody>
</table>

To discard significant within-item differences, a linear mixed effects model was run as above that included a random slope for Item to the random factor Participant (1+Item|Participant). Predictably again, results showed no significant main effect of Group (t(37) = 1.45, p = .155), Mode (t(215) = 0.97, p = .330), or Group*Mode interactions (t(215) = -0.45, p = .555).
F2 (B2). Figure 6 shows the boxplot representing the F2 production of the Spanish front mid vowel /e/ by language dominance and language mode. Both groups presented a somewhat variable range of F2 frequencies, the median frequency being very similar across group and language mode.

To determine whether there were significant differences in the vowel frontness dimension when the target vowel was produced by each language dominance group in the separate language mode conditions, a linear mixed-effects model was run for the F2 data, with Group (2 levels: GAL DOM; SPA DOM) and Mode (2 levels: Monolingual Spanish Mode; Bilingual Galician-Spanish Mode) as fixed factors (including interaction) and Participant as a random factor. The results predictably showed no significant main effect of Group or Mode. In addition, they showed no significant Group*Mode interactions as expected.
Table 6. Results of mixed-effects model for Spanish vowel /e/ Bark-converted second formant (F2)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>12.68185</td>
<td>0.20976</td>
<td>29.87000</td>
<td>60.460</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>Group</td>
<td>0.39134</td>
<td>0.27124</td>
<td>30.07000</td>
<td>1.443</td>
<td>0.159</td>
</tr>
<tr>
<td>Mode</td>
<td>-0.05393</td>
<td>0.12815</td>
<td>218.02000</td>
<td>-0.499</td>
<td>0.618</td>
</tr>
<tr>
<td>Group*Mode</td>
<td>-0.03608</td>
<td>0.16527</td>
<td>218.00000</td>
<td>-0.218</td>
<td>0.827</td>
</tr>
</tbody>
</table>

To discard significant within-item differences had an effect on results, a linear mixed effects model was run as above that included a random slope for Item to the random factor Participant (1+Item|Participant). Predictably again, results showed no significant main effect of Group (t(30) = 1.44, p = .159), Mode (t(218) = -0.50, p = .618), or Group*Mode interactions (t(218) = -0.22, p = .827).

4. Spanish back mid vowel /o/

F1 (B1). Figure 7 shows the boxplot representing the F1 production of the Spanish back mid vowel /o/ by language dominance and language mode. Both groups presented a very steady production across group and language mode.

Figure 7. Boxplot of first formant (F1) Bark-converted distributions for Spanish /o/ by participants’ language dominance and language mode.
To determine whether there were significant differences in the vowel height dimension when the target vowel was produced by each language dominance group in the separate language mode conditions, a linear mixed-effects model was run for the F1 data, with Group (2 levels: GAL DOM; SPA DOM) and Mode (2 levels: Monolingual Spanish Mode; Bilingual Galician-Spanish Mode) as fixed factors (including interaction) and Participant as a random factor. The results predictably showed no significant main effect of Group or Mode. In addition, they showed no significant Group*Mode interactions as expected.

Table 7. Results of mixed-effects model for Spanish vowel /o/ Bark-converted first formant (F1)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.07411</td>
<td>0.16203</td>
<td>31.83</td>
<td>31.317</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>Group</td>
<td>-0.08647</td>
<td>0.20932</td>
<td>31.92</td>
<td>-0.413</td>
<td>0.682</td>
</tr>
<tr>
<td>Mode</td>
<td>0.02069</td>
<td>0.10921</td>
<td>213.13</td>
<td>0.189</td>
<td>0.850</td>
</tr>
<tr>
<td>Group*Mode</td>
<td>-0.05474</td>
<td>0.14188</td>
<td>213.12</td>
<td>-0.456</td>
<td>0.649</td>
</tr>
</tbody>
</table>

To discard significant within-item differences, a linear mixed effects model was run as above that included a random slope for Item to the random factor Participant (1+Item|Participant). Predictably again, results showed no significant main effect of Group \(t(31.6) = -0.39, p = .695\), Mode \(t(187) = 0.22, p = .825\), or Group*Mode interactions \(t(187) = -0.49, p = .627\).

F2 (B2). Figure 8 shows the boxplot representing the F2 production of the Spanish back mid vowel /o/ by language dominance and language mode. Both groups presented a very steady production across group and language mode.
To determine whether there were significant differences in the vowel frontness dimension when the target vowel was produced by each language dominance group in the separate language mode conditions, a linear mixed-effects model was run for the F1 data, with Group (2 levels: GAL DOM; SPA DOM) and Mode (2 levels: Monolingual Spanish Mode; Bilingual Galician-Spanish Mode) as fixed factors (including interaction) and Participant as a random factor. The results predictably showed no significant main effect of Group or Mode. In addition, they showed no significant Group*Mode interactions as expected.

Table 8. Results of mixed-effects model for Spanish vowel /o/ Bark-converted second formant (B2)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>8.64400</td>
<td>0.13103</td>
<td>41.17000</td>
<td>65.971</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>Group</td>
<td>0.04294</td>
<td>0.16938</td>
<td>41.38000</td>
<td>0.254</td>
<td>0.801</td>
</tr>
<tr>
<td>Mode</td>
<td>0.02471</td>
<td>0.12406</td>
<td>213.19000</td>
<td>0.199</td>
<td>0.842</td>
</tr>
<tr>
<td>Group*Mode</td>
<td>0.01225</td>
<td>0.16117</td>
<td>213.17000</td>
<td>0.076</td>
<td>0.939</td>
</tr>
</tbody>
</table>
To discard significant within-item differences had an effect on results, a linear mixed effects model was run as above that included a random slope for Item to the random factor Participant (1+Item|Participant). Predictably again, results showed no significant main effect of Group (t(37.6) = -0.27, p = .785), Mode (t(188) = 0.20, p = .837), or Group*Mode interactions (t(188) = 0.13, p = .899).

**Main findings**

1. **Galician front close/open mid vowel pair /e-ɛ/**

Results of both F1 and F2 analyses show that, contrary to previous research findings, neither the language mode in which the target vowels were produced, nor the participants’ language dominance, affected height or frontness in the production of the target Galician front close mid vowel /e/ and front open mid vowel /ɛ/.

2. **Galician back close/open mid vowel pair /o-ɔ/**

Results of both F1 and F2 analyses show that neither the language mode in which the target vowels were produced, nor the participants’ language dominance, affected height or frontness in the production of the target Galician back close mid vowel /o/ and back close mid vowel /ɔ/.

3. **Spanish front mid vowel /e/**

Results of both F1 and F2 analyses show that as would be expected neither the language mode in which the target vowel was produced, nor the participants’ language dominance had a significant effect on height or frontness in the production of the target Spanish front mid vowel /e/.

4. **Spanish back mid vowel /o/**

Results of both F1 and F2 analyses show that as would be expected neither the language mode in which the target vowel was produced, nor the participants’ language dominance had a significant effect on height or frontness in the production of the target Spanish back mid vowel /o/.
Discussion

This study aimed to investigate the production of the Galician mid vowel close/open contrasts (/e-ɛ/ and /o-ɔ/) and the Spanish mid vowels (/e/, /o/) by two groups of Galician dominant and Spanish dominant proficient sequential bilinguals (n=25). Vowels were generated within monolingual (Galician or Spanish) and code-switched (Bilingual) language mode conditions. The specific goal was to test whether transient effects from cross-linguistic interference (CLI) would influence vowel production values when produced in a code-switched, bilingual language mode, as compared to production in the monolingual mode.

To that end, spoken data were collected during a cued picture-naming task and analysed acoustically to measure the formant values (F1 and F2) of the target vowels when produced by the two groups of speakers in monolingual and bilingual language mode conditions. The formant data’s Bark-converted values were statistically analysed using a linear mixed effects model to investigate differences between groups, language modes, carrier words (items) and vowels, as well as interactions. Models were built for the Galician front mid and back mid vowel pairs separately, and for the Spanish single front mid and back mid vowels separately.

Given their early Galician experience, the Galician dominant group was expected to maintain separate categories in at least the Galician monolingual mode, and possibly be able to maintain categorical production in the bilingual mode condition. The Spanish dominant group was expected to present no significant difference in both the Galician front and back mid vowel pairs in the monolingual condition, although a difference could be present in the back mid vowel pair (Amengual and Chamorro, 2015) in this condition. Presence of this latter difference could have indicated L2 learning effects for the back mid vowel categories in this group of bilinguals. In turn, the Spanish dominant group would be expected to present no significant mid-vowel differences in their Galician L2 vowel productions in the bilingual mode conditions, as a consequence of transient or long term effects (CLI) from the L1 phonological features onto their L2 phonology (e.g. Olson, 2013). For the Spanish vowels, no significant effects should be relevant (however, see Antoniou et al., 2010).

The results revealed no significant effects or interactions overall for any of the vowels in either of the language modes by either group of bilinguals. This is a somewhat unexpected finding, particularly in regards to the Galician dominant group’s mid-vowels, taking into account early experience in this language.

A number of closely related factors will be explored in an attempt to explain these findings.
Overall, early bilinguals are more equipped to learn to discriminate and produce native-like L1 speech (Flege and Mackay, 2004). Jusczyk (1997) identified that, in regular development, a baby’s acquisition of linguistic knowledge relies on a process of distributional analyses of the input s/he receives, which in the case of phonology includes information on the frequency in which acoustic and phonological features cluster together to form phonological regularities. Models of speech perception from the study of second language speech have attempted to explain how these cognitive processes work in relation to cross-linguistic categorisation patterns, relevant to the development of both early and sequential phonological systems (Colantoni, Steele and Escudero, 2015). One such model is the Second Language Perception Model (L2LP; Escudero and Boersma, 2004), which follows the influential Second Language Model’s concept of equivalence classification (SLM; Flege, 1995). Elements of speech will be perceived by the speaker and allocated to existing categories based on previous experience (Flege, 1987), and so the L2LP proposes that this may be modulated by the speaker’s individual ‘perception grammar’ (Colantoni et al., 2015, p.45). A perception grammar is a system that parses and weighs in continuous acoustic input values and maps them onto abstract phonological categories. Importantly for the present study, the L2LP predicts that the speaker’s specific native grammar will determine their ability to perceive new categories later on in life. A lack of mid-vowel differentiation between the two groups of speakers in this study, could then be indicating that both these groups’ perception grammar presents the number of categories of the Spanish mid vowel set. Research has supported that the native (‘skilled’: Pallier et al., 2001) speech perception mechanism focuses on more abstract linguistic information than a non-native one, which offers supports to the theory that difficulties incorporating late categories may be partly attentional (Flege, 1995; Jusczyk, 1997). This raises the question of whether it can be claimed with confidence that the early Galician speakers in the present study had simply not been exposed to enough categorical acoustic data to incorporate the Galician mid vowel categories. Given that Galician and Spanish are languages in long-term contact, and other groups of early Galician speakers in the same geographical area (Vigo) have been shown to present the two sets of mid vowel categories (Amengual and Chamorro, 2015), this seems less likely.

In examining the information from this study’s boxplots (see Results section), it may be noted that the formant values of the individual Galician mid vowel productions present with substantial variability in both modes. Besides potential effects of vowel context, how do we, then, account for this high variability and what else can this say about the bilinguals’ perception grammars?
Variationist accounts of speech perception and production offer a complementary account of how the bilingual’s phonological systems influence each other. In the present study, to account for the bilinguals’ highly variable productions, we must also account for both intralearnner (linguistic) and interlearner variability. In the case of interlearner variability, these would be both internal and external factors to the speaker, so for example, external factors will include the quality and quantity of the input, or the amount of language use; internal factors (individual differences) will include personal qualities such as motivation, and cognitive abilities, such as phonological awareness (see Colantoni, Steele and Escudero, 2015). Ellis (2001) remarked that differing histories of input are what define the differences between speakers, but this input is also enhanced by interpersonal sharing of linguistic experiences, so that variability will veer around patterns of social interaction. For the present study results, this means that the phonological behaviours of individual speakers in the two groups of bilinguals would have been influenced by a host of variable factors influencing these behaviours, some of which would be the result of interspeaker interactions, some others the result of their perception grammars corresponding production, some further still the result of interlearner variability.

Labov (1972) showed that linguistic variability is a systematic process, and that speakers follow socially-ruled linguistic behaviour within selected contexts in order to signify non-linguistic information. Repetitive sociolinguistic patterns of behaviour will eventually propagate a change through the speech community, which will then be reincorporated into the bilinguals’ grammars and so result in diachronic language change (Thomas, 2011). The sociolinguistic landscape of the Galician bilingualism is complex (O’Rourke and Ramallo, 2013; Turell, 2001). It could be argued that the historical context of diglossia (Ferguson, 1959) that for so long positioned the Galician language as the Low prestige variety within the language contact situation, in relation to the High prestige position of the Spanish language, has influenced and perpetuated somewhat contradictory sociophonetic behaviour. Spanish-sounding Galician speech is regarded even by similar groups of speakers as both prestigious (‘normalised Galician’, see Ramallo and Rei-Doval, 2015) and ‘defective’ (O’Rourke and Ramallo, 2015; 2013). This is the case for those bilingual speakers who index themselves as ‘neofalantes’ (‘new speakers’). Even though the label ‘neofalantes’ applies to a specifically-indexed type of speaker, i.e. those with no early linguistic experience in Galician (Tome-Lourido and Evans, 2015), impressionistic observations from the present study seem to point to some early Galician bilinguals reclaiming the neofalantes persona. A strong ideological motivation towards the use of the Galician language is one of the intrapersonal variables that is most often documented amongst neofalantes (O’Rourke and Ramallo, 2015). In some social contexts, this is a mark of
prestige, and bilinguals may index themselves as neofalantes if they perceive their linguistic histories resemble those of others in the category. However, in some cases it is highly likely that early Galician experience is not considered enough for these speakers to feel like they have native Galician speech. For example, some neofalantes may have been raised in a Galician-speaking household but consider themselves L2 Galician speakers because a previous change in language dominance due to the effect of the majority Spanish language. Some of the participants in the present study are early Galician dominant speakers who identify as neofalantes. The relationship of this indexing to long-term contact language history, specifically as found in the Vigo urban areas, may be relevant to the understanding of these speakers’ perception grammars as possibly mirroring a process of convergence of the minority language towards the majority one.

Language contact, substrate effects and prestige

As explained above, variables leading to language change are many and multifactorial. One way to classify linguistic change is to make a distinction between change that is internally motivated and change that is externally motivated (Thomas, 2011). Causes of externally-motivated change include language contact (change spreading from one language to another), and prestige-related changes (change spreading from privileged groups to less privileged groups).

Two languages in long-term contact often develop structural similarities (Bullock and Gerfen, 2004; Chang, 2009). According to Thomason and Kaufman’s (1988) ‘theory of language contact’, one such type of contact is interference, in which the mass acquisition of a second language will cause the affected group to undergo language shift. The most salient effects from this shift are phonetic and phonological changes, which are often perceptually classed as (foreign) accent. As mentioned before, language change is driven by social factors, such as prestige and dominance (Labov, 1972; Backus, 2004). These two points seem to be relevant to the results found in the present study. The Galician open/close mid-vowels seem to have been neutralised into a single allophonic category. This has been partly shown by recent phonetic research findings for the same vowel pairs in Vigo (Amengual and Chamorro, 2015), and other researchers have also been documenting progressive processes of vowel neutralisation, some of which have been claimed are in progress in the Southwesest regions, where Vigo is located (Regueira, 2009). Widespread contact-induced morphological change has been further documented from cartography/dialectology (e.g.Dubert and Sousa, 2016), showing that substrate effects are possibly already be well established at the level of phonology. Phonetic convergence has been found in other long-term contact
languages (for example Mayr et al., 2015 for the Welsh-English pair; Bullock and Gerfen, 2004 for the French-English pair; Recarens and Espinosa, 2003 for the Catalan-Spanish pair).

In respect to prestige-related change, Vidal Figueroa (1997) studied two socially-indexed Galician accents in Vigo, which he named ‘traditional Vigo dialect’ and ‘educated urban Galician’. He found that, while the traditional Vigo dialect group retained a mid-vowel distinction, Galician mid-vowel production was perceptually identical to the Spanish mid-vowels when spoken by members of the educated urban group (the prestigious group). The traditional/urban varieties paradigm adds a further layer to the sociophonetic complexity of the Galician contact-language situation when we consider the sociolinguistic status neofalante groups take within the community (O’Rourke and Ramallo, 2015). Speakers from the traditional Vigo dialect, who retain the mid-vowel distinction, represent the lower prestige, less formally educated population from the surrounding urban centres. Speakers from the educated urban Galician group, who produce Spanish-like mid vowels, are the higher prestige, college educated speakers in more metropolitan urban areas. Galician-specific speech features, sociophonetic behaviour would likely be particularly realised by this group in an attempt to index themselves at once with prestige varieties and ‘traditional’ lower prestige ones, that they attempt to reclaim into higher prestige. As illustrated before, some of the phonological behaviours arising from this have been partly described by Tome-Lourido and Evans (2015). In the Northern regions of Galicia where this study was undertaken, neofalantes have adopted a merged mid-vowel category that is dissimilar from both the Galician and the Spanish mid vowels. It would remain to be explored whether neofalantes could be found as a distinctive social group in the urban Vigo area, or whether matters of indexicality with this group extending to early Galician bilinguals in the South would be enough to propagate future phonological change to this urban centre.

Results from the present study should be taken cautiously. Simonet (2013) found that bilingual language mode effects were likely to prevail once established at an interpersonal level between speakers after brief exchanges. It is possible that the present task design was not fully achieving a threshold level of monolingual activation, and that bilinguals maintained the bilingual language mode on some level during the whole session. Disentangling the effects of long-term bilingualism in the phonological behaviour of contact communities is a multifaceted process that requires experimental design to target highly interconnected (internal/external) psycholinguistic/sociolinguistic variables. Lateral transfer of information between sociolinguistic knowledge and speech perception is a dynamic process that involves parallel cognitive processing between all areas of the cognitive speech perception-production mechanism (Thomas, 2011). Experimental laboratory phonetics offers a chance to control for some of these variables, but it is necessarily limited to practical and resource-related factors.
as well as not being able to account well for naturalistic speech. A different experimental design that would control for language mode differently would enhance the present study’s design by using monolingual speakers to prime language mode (e.g. Mayr et al., 2015; Simonet, 2013).

Two observations from the data are of particular interest: Firstly, the pairs of Galician mid-vowels appear to have been merged into one common phonetic space, reducing the Galician vowel set from 7 to 5; Secondly, speakers’ mid-vowel formant values are highly variable. While these results are intriguing, it remains to be explored whether the vowel productions in the data corpora share similar acoustic value ranges, and where exactly they are positioned in relation to each other in this phonetic space.

In relation to language change, could we predict whether results from the current study may be indicating prestige-induced change or contact-induced change in the Galician mid-vowels? Further study involving a separate group of neofalantes would contribute to the understanding of how pervasive the process of vowel neutralisation may be, and to what extent indexicality may (or may not) be influencing phonological change. Additionally, the present study would have benefited from a speech perception task to complement the data.

This study informs processes related to the developmental sequence of speech acquisition. Bilingual children acquiring language in Galician-dominant families in the Vigo urban area will likely develop the phonological categories described above for this type of bilingualism. A child’s vowel inventory not containing the Galician mid-vowels may not be taken as an indication of vowel error or a language disorder, but should instead be considered to be part of the local accent/phonological patterns of vowel neutralisation in which the child is developing.

**Conclusion**

This study investigated the production of the Galician mid vowels /e- ɛ/ and /o- ɔ/, and the Spanish mid vowels /e/, /o/ in monolingual and code switching modes. Results showed no significant effects or interactions overall for any of the vowels in either of the language modes by either group of bilinguals, indicating they shared comparable underlying categories for these vowels. Interestingly, early Galician speakers with a pattern of Galician dominance did not show categorical differentiation of the Galician close/open mid vowel pairs, possibly indicating they are merging those categories into one. This could support previous research pointing at vowel neutralisation in the Vigo area, but further study is needed. Ultimately, this study contributes to knowledge on the phonological behaviours of bilingual speakers of long-term contact languages.
Reference list


Birdsong, D., Gertken, L.M., & Amengual, M. (2012). Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism. COERLL, University of Texas at Austin [Online: https://sites.la.utexas.edu/bilingual/]


Boersma & Weenink (2017). PRAAT. Software


Grosjean, F. (2012). An attempt to isolate, and then differentiate, transfer and interference. International *Journal of Bilingualism*, 16(1), 11-21


Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36, 3-15


Appendices: Contents

v. Ethics Approval Letter

vi. From Methods: Complementary information
   a. Participant information sheet
   b. Consent form
   c. Participant background questionnaire
   d. Target word lists

vii. From Results: Complementary information
   a. Full data results for linear mixed models with boxplots

viii. Word count
Monday, 21 November 2016

cshs/ethics/approved - interim/

Lopez-Bueno, Laura
BSc (Hons) Speech & Language Therapy
Cardiff School of Health Sciences

Dear Applicant


Ethics Reference Number : 8455

Your ethics application, as shown above, was considered by the Health Care and Food Ethics Panel on 16/11/2016

I am pleased to inform you that your application for ethical approval was APPROVED, subject to the conditions listed below – please read carefully.

Standard Conditions of Approval

- Your Ethics Application has been given a Project Reference number as above. This MUST be quoted on all documentation relating to the project (e.g. consent forms, information sheets), together with the full project title.

- All documents must also have the approved University Logo and the Version number in addition to the reference and project title as above.

- A full Risk Assessment must be undertaken for this proposal, as appropriate, and be made available to the Committee if requested.

- Any changes in connection to the proposal as approved must be referred to the Panel/Committee for consideration without delay quoting your Project Reference Number. Changes to the proposed project may have ethical implications and so must be approved.
• Any untoward incident which occurs in connection with this proposal must be reported back to the Panel/Committee \textit{without delay}.

• If your project involves the use of human samples, your approval is given on the condition that you or your supervisor notify the HTA Designated Individual of your intention to work with such material by completing the form entitled “Notification of Intention to Work with Human Samples”. The form must be submitted to the PD (Sean Duggan), BEFORE any activity on this project is undertaken.

This approval expires on 16/11/2017. Please set a reminder on your Outlook calendar or equivalent if you need to continue beyond this approval date. It is your responsibility to reapply / request extension if necessary.
Yours sincerely

[Signature]

Prof. Arthur Tatham
Chair of Department of Healthcare and Food Ethics Panel
Cardiff School of Health Sciences
Llandaf Campus
Western Avenue, Cardiff CF5 2YB
Tel: 029 2041 7125
E-mail: atatham@cardiffmet.ac.uk

Cc: Mayr, Robert

PLEASE RETAIN THIS LETTER FOR REFERENCE
PARTICIPANT INFORMATION SHEET


Student researcher: Laura López-Bueno
Supervisor: Dr Robert Mayr (rmayr@cardiffmet.ac.uk)

What is the project about?
This study aims to explore how adult bilinguals from Galicia produce Galician and Spanish words, and how their two languages interact. This will add to our understanding of the attention and selection processes during online processing that bilinguals make use of.

Why have I been selected?
You are able to speak Galician and Spanish; you were born in Galicia and have always lived in Galicia; also you received compulsory education in Galician and Spanish.

What will I be asked to do?
With your consent, you will take part in a picture-naming session during which you will be asked to name the words represented in pictures in both Spanish and Galician. The session will take place at a convenient quiet location, either at the researcher’s home, or if you prefer at your home. The session will be audio-recorded.
Can you withdraw from the study?
Yes, your participation in the study is entirely voluntary. You have the right to withdraw from the study at any point during data collection and up to TWO days thereafter.

Will anonymity and confidentiality be guaranteed?
Yes, strict anonymity and confidentiality will be ensured. Any reference to you will be made anonymous, making it impossible to determine your identity. The recordings will be stored on a password-protected computer, with access to the data restricted to the researcher and the supervisor working on the project. The data will be destroyed once the project has been completed unless you agree for the data to be retained for research purposes.

If you have any further questions, please do not hesitate to contact the student’s supervisor Dr Robert Mayr (rmayr@cardiffmet.ac.uk).

Many thanks.

Laura López-Bueno
ii.b. Participant consent form

PARTICIPANT CONSENT FORM

Ethics Reference Number: 8455

Title of Project: Code Switching in Contact Languages: An exploration of language mode effects in Spanish-Galician bilingual speakers.

Name of researcher: Laura López-Bueno

Name and contact details of supervisor: Dr Robert Mayr (mawy@cardiffnet.ac.uk)

Participant to complete this section: Please initial each box.

1. I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

2. I understand that my participation is voluntary and that I am free to withdraw from the study at any time during data collection, and up to TWO days after data collection.

3. I agree to take part in the above study.

4. I agree for the data collected to be retained for research purposes.

PRINT NAME ___________________________ Date ___________________________

Signature of Participant ___________________________ ___________________________

One copy for the participant; one for the researcher
ii.c. Participant background questionnaire

Ethics reference number: XXX

Participant Background Questionnaire


Student researcher: Laura López-Bueno

Supervisor: Dr Robert Mayr (mayr@cardiffmet.ac.uk)

I. PERSONAL DETAILS

1. Name: ____________________________

2. Email: ____________________________

3. Age: ____________________________

4. Place where you were born: ____________________________

5. What do you do for a living? ____________________________

6. What’s your highest level of education? (Please tick)

<table>
<thead>
<tr>
<th>Primary Schooling</th>
<th>[ ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Secondary Schooling</td>
<td>[ ]</td>
</tr>
<tr>
<td>Higher Education</td>
<td>[ ]</td>
</tr>
<tr>
<td>Master’s degree</td>
<td>[ ]</td>
</tr>
<tr>
<td>Professional/Technical degree</td>
<td>[ ]</td>
</tr>
</tbody>
</table>

Are you currently a student? YES/NO (please circle)

What degree(s)/Master's? ____________________________

What professional/technical studies? ____________________________
II. PLACES WHERE YOU LIVED

7. Have you ever lived away from Galicia for longer than one month at any point in your life? (excluding holidays, but including away-stays for professional reasons, courses, etc., if the place where you lived became your place of residence albeit temporarily)
   No  ➔ Jump to question 11.
   Yes  ➔ Go to the next question.

8. Where away from Galicia did you live?

9. What age(s) were you?

10. For how long?

11. Have you ever lived away from Vigo for longer than one month at any point in your life? (excluding holidays, but including away-stays for professional reasons, courses, etc., if the place where you lived became your place of residence albeit temporarily)
   No  ➔ Jump to question 14.
   Yes  ➔ Go to the next question.

12. Where else did you live in Galicia?

13. How long have you been living in Vigo?

III. LANGUAGE USE

14. What age were you when you learnt Galician? How did you learn it?

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

16. What language do you use the most?

17. Has this always been your main language?  YES / NO
18. If you answered **NO** to the previous question, when/what age did you start speaking this language as your main language? 

19. Why did you switch languages? 

20. Where are/were your parents from (or the people who raised you), and what is their main language? 

<table>
<thead>
<tr>
<th>Person (mother, father, grandmother, grandfather...)</th>
<th>Place of origin</th>
<th>Main language</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

21. Have they always spoken these languages? 

22. List any languages you speak other than Galician and Spanish, indicate your competence in them, and how often and for what purpose you use them. 

<table>
<thead>
<tr>
<th>Language</th>
<th>Competence</th>
<th>Frequency and purpose of use</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
23. The following table asks you to think about how much you use Galician with certain people and in certain situations. Next to each person or situation, estimate how often you use Galician and tick the closest option that applies.

Sometimes you'll be asked to specify how often you do a particular activity or see a certain person (you may respond, for example, ONCE A WEEK, or EVERY WEEKEND).

<table>
<thead>
<tr>
<th>Situation</th>
<th>Never</th>
<th>Rarely</th>
<th>About half the time</th>
<th>Often</th>
<th>Always</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brothers/Sisters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grandfather</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grandmother</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teachers/tutors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At work</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At college/university</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With your partner</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often do you see each other?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With your best friend</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often do you see each other?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With your work colleagues</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often do you work?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With other people at work (e.g. clients)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With your college/uni peers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often do you go to college/uni?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With other people at college/uni</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Think about any other things you do where you might (or might not) use Galician and which are not listed above. **Write how often you do this activity in brackets.**

<table>
<thead>
<tr>
<th>Situation</th>
<th>Never</th>
<th>Rarely</th>
<th>About half the time</th>
<th>Often</th>
<th>Always</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example: With the other people I’m in a band with (ONCE A WEEK).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

24. **Any other comments?**

*Thank you very much for completing this survey.*
ii.d. Full list of Target words included in the trials

**Token words front vowels**

<table>
<thead>
<tr>
<th>GALICIAN /ə/</th>
<th>GALICIAN /ɛ/</th>
<th>SPANISH /e/</th>
</tr>
</thead>
<tbody>
<tr>
<td>pelo</td>
<td>ˈpelo</td>
<td>regra</td>
</tr>
<tr>
<td>neno</td>
<td>ˈnenɔ</td>
<td>vela</td>
</tr>
<tr>
<td>dedo</td>
<td>ˈdedɔ</td>
<td>vello</td>
</tr>
<tr>
<td>cesta</td>
<td>ˈθɛstɔ</td>
<td>perna</td>
</tr>
<tr>
<td>cebra</td>
<td>ˈθɛβɾa</td>
<td>césar</td>
</tr>
<tr>
<td>testo</td>
<td>ˈtestɔ</td>
<td>serpe</td>
</tr>
<tr>
<td>seto</td>
<td>ˈsetɔ</td>
<td>sete</td>
</tr>
<tr>
<td>cepo</td>
<td>ˈθɛpɔ</td>
<td>tecla</td>
</tr>
</tbody>
</table>

**Token words back vowels**

<table>
<thead>
<tr>
<th>GALICIAN /o/</th>
<th>GALICIAN /ɔ/</th>
<th>SPANISH /o/</th>
</tr>
</thead>
<tbody>
<tr>
<td>fonte</td>
<td>ˈfontɛ</td>
<td>foca</td>
</tr>
<tr>
<td>goma</td>
<td>ˈgɔmɔ</td>
<td>copa</td>
</tr>
<tr>
<td>sobre</td>
<td>ˈsoβɾɛ</td>
<td>sota</td>
</tr>
<tr>
<td>oso</td>
<td>ˈosɔ</td>
<td>roda</td>
</tr>
<tr>
<td>mono</td>
<td>ˈmɔnɔ</td>
<td>nove</td>
</tr>
<tr>
<td>cono</td>
<td>ˈkɔnɔ</td>
<td>porta</td>
</tr>
<tr>
<td>polo</td>
<td>ˈpɔlɔ</td>
<td>norte</td>
</tr>
<tr>
<td>lobo</td>
<td>ˈloβɔ</td>
<td>óso</td>
</tr>
</tbody>
</table>
iii.a. Full data results for linear mixed models and additional boxplots

1.a. Results of mixed-effects model for Galician vowel pair /e-ɛ/ Bark-converted first formant (F1)

```r
> gale.model.lme(lm(Block|Group+Participant|Participant)+Group*Vowel+Group*Vowel+Group*Vowel+Vowel+Node|data=gale,REML=FALSE)
> summary(gale.mixed)
```

Linear mixed model fit by maximum likelihood

| t-test use Satterthwaite approximations to degrees of freedom ['lmerMod'] |
|-----------------------------|-----------------------------|
| Formula: B1 ~ Group + Node + Vowel + (1 + Vowel | Participant) + Group * Vowel + Group + Node + Vowel + Node + Group + Vowel + Node |
| Data: gale |

<table>
<thead>
<tr>
<th>AIC</th>
<th>BIC</th>
<th>logLik deviance df.residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>313.7</td>
<td>351.6</td>
<td>-156.3 192.7</td>
</tr>
</tbody>
</table>

Scaled residuals:

<table>
<thead>
<tr>
<th>Min</th>
<th>1Q</th>
<th>Median</th>
<th>3Q</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2.5000</td>
<td>-0.5970</td>
<td>-0.1142</td>
<td>0.3465</td>
<td>4.6256</td>
</tr>
</tbody>
</table>

Random effects:

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance</th>
<th>Std.Dev.</th>
<th>Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>Intercept</td>
<td>0.128351</td>
<td>0.3534</td>
<td></td>
</tr>
<tr>
<td>Vowel</td>
<td>0.0001315</td>
<td>0.0363</td>
<td>-1.08</td>
<td></td>
</tr>
<tr>
<td>Node</td>
<td>0.2741595</td>
<td>0.5238</td>
<td>-1.00</td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>0.25625</td>
<td>0.5062</td>
<td>-1.01</td>
<td></td>
</tr>
<tr>
<td>Group:Vowel</td>
<td>-0.18833</td>
<td>0.3437</td>
<td>0.7984</td>
<td>0.4022</td>
</tr>
</tbody>
</table>

Fixed effects:

| Estimate | Std. Error | df | t value | Pr(>|t|) |
|----------|------------|----|---------|----------|
| Intercept| 0.103551   | 395.00000 | 33.392 | 2e-16 ***|
| Group    | 0.0134951  | 0.01355 | 1.017  | 0.909 |
| Node     | 0.010647   | 0.01064 | 1.000  | 0.520 |
| Vowel    | 0.010620   | 0.01062 | 1.000  | 0.520 |
| Node:Vowel| 0.013055 | 0.01305 | 1.000  | 0.520 |
| Group:Node:Vowel | 0.010870 | 0.01087 | 1.000  | 0.520 |

Correlation of Fixed Effects:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intr</td>
<td>1.000</td>
<td>-1.00</td>
<td>0.798</td>
<td>-1.00</td>
<td>0.520</td>
<td>-1.01</td>
<td>0.798</td>
<td>-1.01</td>
<td>-1.01</td>
<td>0.798</td>
<td>-1.01</td>
</tr>
</tbody>
</table>

1.b. Results of mixed-effects model for Galician vowel pair /e-ɛ/ Bark-converted second formant (F2)

```r
> gale.model.lme(lm(Block|Group+Participant|Participant)+Group*Vowel+Group*Vowel+Group*Vowel+Vowel+Node|data=gale,REML=FALSE)
> summary(gale.mixed)
```

Linear mixed model fit by maximum likelihood

| t-test use Satterthwaite approximations to degrees of freedom ['lmerMod'] |
|-----------------------------|-----------------------------|
| Formula: B2 ~ Group + Node + Vowel + (1 + Vowel | Participant) + Group * Vowel + Group + Node + Vowel + Node + Group + Vowel + Node |
| Data: gale |

<table>
<thead>
<tr>
<th>AIC</th>
<th>BIC</th>
<th>logLik deviance df.residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>313.7</td>
<td>351.6</td>
<td>-156.3 192.7</td>
</tr>
</tbody>
</table>

Scaled residuals:

<table>
<thead>
<tr>
<th>Min</th>
<th>1Q</th>
<th>Median</th>
<th>3Q</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2.4122</td>
<td>-0.2992</td>
<td>0.0192</td>
<td>0.4020</td>
<td>2.7724</td>
</tr>
</tbody>
</table>

Random effects:

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance</th>
<th>Std.Dev.</th>
<th>Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>Intercept</td>
<td>0.3081</td>
<td>0.5560</td>
<td>-0.76</td>
</tr>
<tr>
<td>Vowel</td>
<td>0.5992</td>
<td>0.7760</td>
<td>-0.15</td>
<td></td>
</tr>
<tr>
<td>Node</td>
<td>0.25625</td>
<td>0.5062</td>
<td>-1.00</td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>0.3255</td>
<td>0.5704</td>
<td>-1.00</td>
<td></td>
</tr>
<tr>
<td>Group:Vowel</td>
<td>0.5120</td>
<td>0.7154</td>
<td>-1.00</td>
<td></td>
</tr>
<tr>
<td>Group:Node</td>
<td>0.5500</td>
<td>0.7408</td>
<td>-1.00</td>
<td></td>
</tr>
<tr>
<td>Group:Node:Vowel</td>
<td>0.5120</td>
<td>0.7154</td>
<td>-1.00</td>
<td></td>
</tr>
</tbody>
</table>

Fixed effects:

| Estimate | Std. Error | df | t value | Pr(>|t|) |
|----------|------------|----|---------|----------|
| Intercept| 0.1034551  | 395.00000 | 33.392 | 2e-16 ***|
| Group    | 0.0134951  | 0.01355 | 1.017  | 0.909 |
| Node     | 0.010647   | 0.01064 | 1.000  | 0.520 |
| Vowel    | 0.010620   | 0.01062 | 1.000  | 0.520 |
| Node:Vowel| 0.013055 | 0.01305 | 1.000  | 0.520 |
| Group:Node:Vowel | 0.010870 | 0.01087 | 1.000  | 0.520 |

Correlation of Fixed Effects:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intr</td>
<td>1.000</td>
<td>-1.00</td>
<td>0.798</td>
<td>-1.00</td>
<td>0.520</td>
<td>-1.01</td>
<td>0.798</td>
<td>-1.01</td>
<td>-1.01</td>
<td>0.798</td>
<td>-1.01</td>
</tr>
</tbody>
</table>
1.c. Results of mixed-effects model with *Item* for Galician vowel pair /e-ɛ/ Bark-converted first formant (F1)

\[ \text{formula} \]

\[ \text{item} \]

\[ \text{formula} \]

\[ \text{item} \]

1.d. Results of mixed-effects model with *Item* for Galician vowel pair /e-ɛ/ Bark-converted second formant (F2)

\[ \text{formula} \]

\[ \text{item} \]
1.c/d.a. Boxplots of first and second formants (F1 and F2) Bark-converted distributions for Galician /e/ and /ɛ/ by target word (Item), language dominance (Group) and language mode (Mode).

Mode Item Group

![Boxplots of first formant (F1) for Galician /e/](image1)

![Boxplots of second formant (F2) for Galician /e/](image2)
2.a. Results of mixed-effects model for Galician vowel pair /o-ɔ/ Bark-converted first formant (F1)

```
galo.model<-lmer(B1~Group+Node+Vowel+1+Vowel|Participant)+Group+Vowel+Node+Group+Vowel+Node+Node, data=galo,REML=FALSE)
> summary(galo.model)
```

Linear mixed model fit by maximum likelihood

- t-tests use Satterthwaite approximations to degrees of freedom [lmerMod]

```
Formula: B1 ~ Group + Node + Vowel + 1 + Vowel | Participant ) + Group + Vowel + Node + Vowel + Node + Group + Vowel + Node
Data: galo

AIC BIC logLik deviance df.resid
981.4 1011.8 473.7 987.4 481

Scaled residuals:
Min 1Q Median 3Q Max
-5.9502 -0.4660 -0.0520 0.4890 5.6099

Random effects:
Groups Name Variance Std.Dev. Corr
Participant (Intercept) 0.1840 0.4295 Vowel 0.8347 0.9486
Residual 0.5047 0.7096
Number of obs: 493; groups: Participant, 25

Fixed effects:
```
Estimate Std. Error t value Pr(>|t|)
(Intercept) 4.73360 0.15110 31.402 2e-16 ***
Group 0.07045 0.01586 4.411 2e-04 ***
Node 0.00256 0.12121 0.021 0.983
Vowel 0.18055 0.13422 13.408 2e-16 ***
Group:Vowel 0.21780 0.19738 1.102 0.275
Group:Node 0.00219 0.12540 0.017 0.984
Node:Vowel 0.10851 0.13081 0.829 0.407
Group:Node:Vowel 0.18335 0.12841 45.450 2e-06 ***
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Correlation of Fixed Effects:
```
Group -0.724
Node -0.401 0.311
Vowel -0.504 0.309 0.452
Group:Vowel 0.309 0.504 -0.340 0.773
Group:Node 0.338 -0.401 0.774 0.426 0.309 0.452
Node:Vowel 0.205 -0.221 -0.711 -0.035 0.491 0.358
Group:Node:Vowel 0.212 0.235 0.530 0.491 0.358 -0.711 -0.452
```

2.b. Results of mixed-effects model for Galician vowel pair /o-ɔ/ Bark-converted second formant (F2)

```
galo.model<-lmer(B2~Group+Node+Vowel+1+Vowel|Participant)+Group+Vowel+Node+Group+Vowel+Node+Node, data=galo,REML=FALSE)
> summary(galo.model)
```

Linear mixed model fit by maximum likelihood

- t-tests use Satterthwaite approximations to degrees of freedom [lmerMod]

```
Formula: B2 ~ Group + Node + Vowel + 1 + Vowel | Participant ) + Group + Vowel + Node + Vowel + Node + Group + Vowel + Node
Data: galo

AIC BIC logLik deviance df.resid
990.0 995.4 -401.0 922.0 481

Scaled residuals:
Min 1Q Median 3Q Max
-6.0843 -0.6934 -0.0080 0.5935 7.0768

Random effects:
Groups Name Variance Std.Dev. Corr
Participant (Intercept) 0.05569 0.2358 Vowel 0.81236 0.9050 1.00
Residual 0.34773 0.5900
Number of obs: 493; groups: Participant, 25

Fixed effects:
```
Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.08775 0.01104 7.9928 7e-15 ***
Group 0.02925 0.01427 2.065 0.039 **
Node 0.18056 0.11920 1.517 0.130
Vowel 0.05556 0.11275 0.491 0.621
Group:Vowel 0.19218 0.15494 0.126 0.899
Group:Node 0.01755 0.13549 0.130 0.898
Node:Vowel 0.24493 0.18771 1.294 0.199
Group:Node:Vowel 0.22544 0.21061 1.067 0.282
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Correlation of Fixed Effects:
```
Group -0.774
Node -0.548 0.410
Vowel -0.236 0.269 0.402
Group:Vowel 0.200 -0.337 -0.373 -0.773
Group:Node 0.417 0.249 -0.774 -0.273 0.462
Node:Vowel 0.383 -0.297 -0.711 -0.078 0.524 0.550
Group:Node:Vowel -0.297 0.504 0.550 0.524 -0.079 -0.711 -0.774
```
2.c. Results of mixed-effects model with Item for Galician vowel pair /o-ɔ/ Bark-converted first formant (F1)

```r
gala.mixed <- lmer(Inv.error ~ (1|Item)(1|Participant) + (1|Item)(1|Participant) + (1|Item)(1|Participant) + (1|Item)(1|Participant) + (1|Item)(1|Participant), data=gala)
```

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.0461</td>
<td>0.1516</td>
<td>-0.302</td>
<td>0.766</td>
</tr>
<tr>
<td>Item</td>
<td>0.1565</td>
<td>0.0588</td>
<td>2.665</td>
<td>0.008</td>
</tr>
<tr>
<td>Vowel</td>
<td>0.1261</td>
<td>0.0612</td>
<td>2.068</td>
<td>0.040</td>
</tr>
</tbody>
</table>

Random effects:
- Groups: Item, Participant

Covariance of Fixed Effects:
```
   Item  Vowel
Item  0.1565
Vowel 0.0588 0.1464
```

2.d. Results of mixed-effects model with Item for Galician vowel pair /o-ɔ/ Bark-converted second formant (F2)

```r
gala.mixed <- lmer(Inv.error ~ (1|Item)(1|Participant) + (1|Item)(1|Participant) + (1|Item)(1|Participant) + (1|Item)(1|Participant) + (1|Item)(1|Participant), data=gala)
```

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.0461</td>
<td>0.1516</td>
<td>-0.302</td>
<td>0.766</td>
</tr>
<tr>
<td>Item</td>
<td>0.1565</td>
<td>0.0588</td>
<td>2.665</td>
<td>0.008</td>
</tr>
<tr>
<td>Vowel</td>
<td>0.1261</td>
<td>0.0612</td>
<td>2.068</td>
<td>0.040</td>
</tr>
</tbody>
</table>

Random effects:
- Groups: Item, Participant

Covariance of Fixed Effects:
```
   Item  Vowel
Item  0.1565
Vowel 0.0588 0.1464
```

```
2.c/d.a. Boxplots of first and second formants (F1 and F2) Bark-converted distributions for Galician /o/ and /ɔ/ by target word (Item), language dominance (Group) and language mode (Mode).

Mode Item Group

![Boxplot of Frequency F1 for Galician /o/](image1)

![Boxplot of Frequency F2 for Galician /o/](image2)
3.a. Results of mixed-effects model for Spanish vowel /e/ Bark-converted first formant (F1)

> spm.model1<lm(B1~Group+Node+(1|Participant)+Group+Node, data=spm, REML=FALSE)
> summary(spm.model1)
Linear mixed model fit by maximum likelihood
  t-tests use Satterthwaite approximations to degrees of freedom 'lmerMod'
Formula: B1 ~ Group + Node + (1 | Participant) + Group + Node
Data: spm

          Estimate Std. Error t value  Pr(>|t|)    
(Intercept) 4.65469 0.125653  36.8000  <2e-16 ***
Group       0.29093 0.166550   1.7405    0.0898 .
Node        0.09794 0.103536   0.9452    0.3483
Group:Node -0.06646 0.123551  -0.5363    0.5930

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
  (Intr) Group Node
Group -0.772
Node -0.359 0.302
Group:Node 0.302 -0.395 -0.775

3.b. Results of mixed-effects model for Spanish vowel /e/ Bark-converted second formant (F2)

> spm.model2<lm(B2~Group+Node+(1|Participant)+Group+Node, data=spm, REML=FALSE)
> summary(spm.model2)
Linear mixed model fit by maximum likelihood
  t-tests use Satterthwaite approximations to degrees of freedom 'lmerMod'
Formula: B2 ~ Group + Node + (1 | Participant) + Group + Node
Data: spm

          Estimate Std. Error t value  Pr(>|t|)    
(Intercept) 5.58526 0.225072  24.7300  <2e-16 ***
Group       0.39334 0.271240   1.4430    0.1509
Node        0.06393 0.128150   0.4999    0.6216
Group:Node -0.06306 0.165270  -0.3837    0.7027

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
  (Intr) Group Node
Group -0.773
Node -0.259 0.228
Group:Node 0.229 -0.300 -0.775
3.c. Results of mixed-effects model with Item for Spanish vowel /e/ Bark-converted first formant (F1)

```r
> spae.model.1er(B1=Group+Mode+(1|Item|Participant)+Group*Mode+Group*Item+Group*Item*Mode,data=spae,REML=FALSE)
> summary(spaе.model)
```

<table>
<thead>
<tr>
<th>AIC</th>
<th>BIC</th>
<th>loglik</th>
<th>deviance</th>
<th>df.resid</th>
</tr>
</thead>
<tbody>
<tr>
<td>415.0</td>
<td>450.9</td>
<td>-195.5</td>
<td>391.9</td>
<td>231</td>
</tr>
</tbody>
</table>

Scaled residuals:
- Min: -2.0488
- Q1: -0.9992
- Median: 0.4293
- Q3: 0.4293
- Max: 4.2970

Random effects:
- Groups Name: Variance Std.Dev. Corr
- Participant (Intercept): 0.677467 0.2603
- Item: 0.044105 0.2051 1.00
- Residual: 0.243859 0.4943

Number of obs: 243, groups: Participant, 25

Fixed effects:
- Estimate Std. Error t value Pr(>|t|)
- (Intercept) 6.00537 1.42929 4.217 2e-16 ***
- Group 0.10955 0.19360 0.568 0.570
- Mode 0.00055 0.17198 0.032 0.975
- Item 0.03050 0.05004 0.605 0.548
- Group:Mode 0.01055 0.12211 0.174 0.863
- Group:Item 0.05000 0.06821 0.736 0.465
- Mode:Item 0.00710 0.01792 0.394 0.698
- Group:Mode:Item 0.00200 0.00929 0.214 0.832
- Residual 0.03050 0.05004 0.605 0.548

Correlation of Fixed Effects:
- (Intr) Group Mode Item Grp:Mo Grp:It Md:Itm
- Group -0.773
- Mode -0.578 0.469
- Item 0.514 0.405 -0.561
- Group:Mode 0.441 0.572 0.374 0.485
- Group:Item 0.398 0.516 0.417 <0.001 0.502
- Mode:Item 0.454 0.591 0.395 0.262 0.511
- Group:Mode:Item 0.351 0.461 0.627 0.520 0.812 0.685 0.774

3.d. Results of mixed-effects model with Item for Spanish vowel /e/ Bark-converted second formant (F2)

```r
> spae.model.1er(B2=Group+Mode+(1|Item|Participant)+Group*Mode+Group*Item+Group*Item*Mode,data=spae,REML=FALSE)
> summary(spaе.model)
```

<table>
<thead>
<tr>
<th>AIC</th>
<th>BIC</th>
<th>loglik</th>
<th>deviance</th>
<th>df.resid</th>
</tr>
</thead>
<tbody>
<tr>
<td>546.4</td>
<td>580.3</td>
<td>-268.2</td>
<td>520.4</td>
<td>231</td>
</tr>
</tbody>
</table>

Scaled residuals:
- Min: -4.4556
- Q1: -0.3288
- Median: 0.4820
- Q3: 0.4820
- Max: 5.0117

Random effects:
- Groups Name: Variance Std.Dev. Corr
- Participant (Intercept): 0.083=0.01 0.00139
- Item: 2.138=0.01 0.1462 1.00
- Residual: 2.924=0.01 0.17225

Number of obs: 243, groups: Participant, 25

Fixed effects:
- Estimate Std. Error t value Pr(>|t|)
- (Intercept) 17.79051 0.24417 73.5200 9e-16 ***
- Group 0.25292 0.32594 0.778 0.442
- Mode -0.01003 0.21306 0.049 0.961
- Item -0.02250 0.03722 0.910 0.364
- Group:Mode -0.03232 0.28178 0.116 0.907
- Group:Item 0.07266 0.05227 1.394 0.161
- Mode:Item -0.03067 0.05120 0.608 0.548
- Group:Mode:Item -0.01314 0.11785 0.090 0.924
- Residual 0.03050 0.05004 0.605 0.548

Correlation of Fixed Effects:
- (Intr) Group Mode Item Grp:Mo Grp:It Md:Itm
- Group -0.774
- Mode -0.642 0.263
- Item 0.500 0.393 0.375
- Group:Mode 0.363 0.445 -0.747 -0.445
- Group:Item 0.388 0.568 0.347 -0.760 0.578
- Mode:Item 0.353 0.273 0.111 0.688 0.532
- Group:Mode:Item -0.273 0.350 0.627 0.532 -0.812 -0.700 -0.774

10
3.c/d.a. Boxplots of first and second formants (F1 and F2) Bark-converted distributions for Spanish /e/ by target word (Item), language dominance (Group) and language mode (Mode).

Mode Item Group

Spanish /e/

![Boxplot showing distribution of F1 and F2 formants for Spanish /e/](image)

![Boxplot showing distribution of F1 and F2 formants for Spanish /e/](image)
4.a. Results of mixed-effects model for Spanish vowel /o/ Bark-converted first formant (F1)

```r
> speco.model <- lmer(res ~ Group + Node + (1 | Participant) + Group + Node, data = speco, REML = FALSE)
> summary(speco.model)
```

Random effects:
- **Groups**: Name Variance Std.Dev.
  - Participant (Intercept) 0.0392 0.198
  - Residual 0.3577
  - Number of obs: 238, groups: Participant, 25

Fixed effects:
- Estimate  Std. Error df  t value Pr(>|t|)
  - (Intercept) 5.07421  0.16209 12.02000  31.317   <2e-16
  - Group -0.20647  0.09324 11.92000  -0.219 0.826
  - Node 0.02089  0.09221 213.10000  0.229 0.820
  - Group:Node -0.00475  0.04706 213.10000  -0.040 0.968
- **Signif. codes:** 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
- (Intr) Group Node
- Group 0.774
- Node 0.340 0.264
- Group:Node 0.262 -0.340 -0.770

4.b. Results of mixed-effects model for Spanish vowel /o/ Bark-converted first formant (F1)

```r
> speco.model <- lmer(res ~ Group + Node + (1 | Participant) + Group + Node, data = speco, REML = FALSE)
> summary(speco.model)
```

Random effects:
- **Groups**: Name Variance Std.Dev.
  - Participant (Intercept) 0.0379 0.194
  - Residual 0.3729 0.6107
  - Number of obs: 238, groups: Participant, 25

Fixed effects:
- Estimate  Std. Error df  t value Pr(>|t|)
  - (Intercept) 5.04800  0.15160 41.17000  33.676   <2e-16
  - Group 0.04254  0.03936 41.17000  1.080 0.281
  - Node 0.02471  0.02465 213.10000  0.999 0.320
  - Group:Node 0.01225  0.01271 213.17000  0.976 0.330
- **Signif. codes:** 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
- (Intr) Group Node
- Group 0.774
- Node 0.479 0.370
- Group:Node 0.386 -0.477 -0.770
```
4.c. Results of mixed-effects model with Item for Spanish vowel /o/ Bark-converted first formant (F1)

```r
> spaomodel1 <- glmer(B ~ Group + Item + (1 + Item | Participant) + Group + Item, data = spaor, REML = FALSE)
```

```r
> summary(spaomodel1)
```

Linear mixed model fit by maximum likelihood

T-tests use Satterthwaite approximations to degrees of freedom (lmerMod)

Formula: B ~ Group + Item + (1 + Item | Participant) + Group + Item

Data: spaor

<table>
<thead>
<tr>
<th>AIC</th>
<th>BIC</th>
<th>logLik deviance df.resid</th>
</tr>
</thead>
<tbody>
<tr>
<td>440.0</td>
<td>490.4</td>
<td>-222.0 424.0 118</td>
</tr>
</tbody>
</table>

Scaled residuals:

<table>
<thead>
<tr>
<th>Min</th>
<th>Q1</th>
<th>Med</th>
<th>Q3</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3.1285</td>
<td>-0.4808</td>
<td>0.8256</td>
<td>4.2390</td>
<td></td>
</tr>
</tbody>
</table>

Random effects:

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance Std.Dev. Corr</th>
<th>Participant (Intercept)</th>
<th>0.21480 0.46160</th>
<th>Item</th>
<th>0.00232 0.84170</th>
<th>0.22</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Residual</td>
<td>0.27083 0.52550</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Number of obs: 238, groups: Participant, 25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fixed effects:

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|---------|
| Group    | -0.12717   | 0.25214 | 0.504   | 0.607 |
| Item     | 0.23661    | 0.18266 | 1.297   | 0.200 |
| Item     | 0.00000    | 0.00000 | 0.000   | 1.000 |
| Item     | 0.00000    | 0.00000 | 0.000   | 1.000 |
| Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 |

Correlation of Fixed Effects:

<table>
<thead>
<tr>
<th>(Intr)</th>
<th>Group</th>
<th>Mode</th>
<th>Item</th>
<th>Gp:hd</th>
<th>Gp:it</th>
<th>M:it</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>-0.756</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>-0.646</td>
<td>0.365</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>-0.659</td>
<td>0.435</td>
<td>0.552</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>0.262</td>
<td>-0.469</td>
<td>-0.733</td>
<td>-0.842</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>0.695</td>
<td>-0.397</td>
<td>-0.811</td>
<td>-0.688</td>
<td>0.426</td>
<td>0.532</td>
</tr>
<tr>
<td>Item</td>
<td>-0.295</td>
<td>0.381</td>
<td>0.624</td>
<td>0.530</td>
<td>-0.689</td>
<td>-0.664</td>
</tr>
</tbody>
</table>

4.c. Results of mixed-effects model with Item for Spanish vowel /o/ Bark-converted second formant (F2)

```r
> spaomodel2 <- glmer(B ~ Group + Item + (1 + Item | Participant) + Group + Item, data = spaor, REML = FALSE)
```

```r
> summary(spaomodel2)
```

Linear mixed model fit by maximum likelihood

T-tests use Satterthwaite approximations to degrees of freedom (lmerMod)

Formula: B ~ Group + Item + (1 + Item | Participant) + Group + Item

Data: spaor

<table>
<thead>
<tr>
<th>AIC</th>
<th>BIC</th>
<th>logLik deviance df.resid</th>
</tr>
</thead>
<tbody>
<tr>
<td>428.5</td>
<td>488.9</td>
<td>-219.2 418.4 226</td>
</tr>
</tbody>
</table>

Scaled residuals:

<table>
<thead>
<tr>
<th>Min</th>
<th>Q1</th>
<th>Med</th>
<th>Q3</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2.42287</td>
<td>-0.08921</td>
<td>-0.87178</td>
<td>0.01080</td>
<td>0.00000</td>
</tr>
</tbody>
</table>

Random effects:

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance Std.Dev. Corr</th>
<th>Participant (Intercept)</th>
<th>0.38865 0.87154</th>
<th>Item</th>
<th>0.00217 0.50178</th>
<th>1.00</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Residual</td>
<td>0.22052 0.47296</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Number of obs: 238, groups: Participant, 25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fixed effects:

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|---------|
| Group    | 0.18562    | 0.17088 | 77.56000 | 0.681 | 0.500 |
| Mode     | 0.00174    | 0.00164 | 212.90000 | 0.497 | 0.620 |
| Item     | 0.06478    | 0.05262 | 122.09000 | 0.947 | 0.358 |
| Item     | 0.09935    | 0.06811 | 101.12000 | 0.434 | 0.665 |
| Mode:Item| -0.00331   | 0.00023 | 212.26000 | -0.496 | 0.621 |
| Mode:Item| 0.00784    | 0.00860 | 213.14000 | 0.795 | 0.427 |
| Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 |

Correlation of Fixed Effects:

<table>
<thead>
<tr>
<th>(Intr)</th>
<th>Group</th>
<th>Mode</th>
<th>Item</th>
<th>Gp:hd</th>
<th>Gp:it</th>
<th>M:it</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>-0.794</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>-0.625</td>
<td>0.485</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>-0.487</td>
<td>0.377</td>
<td>-0.552</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>0.483</td>
<td>-0.624</td>
<td>-0.733</td>
<td>-0.811</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>0.376</td>
<td>-0.488</td>
<td>-0.011</td>
<td>-0.733</td>
<td>0.530</td>
<td></td>
</tr>
<tr>
<td>Mode:Item</td>
<td>0.511</td>
<td>-0.395</td>
<td>-0.606</td>
<td>0.626</td>
<td>0.513</td>
<td></td>
</tr>
<tr>
<td>Mode:Item</td>
<td>0.200</td>
<td>0.500</td>
<td>0.624</td>
<td>0.512</td>
<td>-0.009</td>
<td>-0.660</td>
</tr>
</tbody>
</table>
4.c/d.a. Boxplots of first and second formants (F1 and F2) Bark-converted distributions for Spanish /o/ by target word (Item), language dominance (Group) and language mode (Mode).

Mode Item Group

Spanish /o/

Monolingual Mode

Bilingual Mode

Frequency F1

Frequency F2

0 Spanish Dominant

1 Galician Dominant
Word Count

Introduction: 89
Literature review: 2310
Methods: 2451
Results: 2554
Discussion: 2585

Total: 9989 Words